

OTHER INFORMATION: n=a, g, c or t

Seq ID no. 129

Query Match 6.5%; Score 35.8; DB 4; Length 3001;  
Best Local Similarity 54.1%; Pred. No. 0.9;  
Matches 73; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 419 TGAATGCAACATTTTCTTAATTAATCTGAACCTGACGATGACCTGAAAAATTAGCTGA 478  
DB 2385 TGCAGATATCATTTTCCCGAGTCTGGAAGCTAGAGATCAAGAAAAACAACAAAA 2326  
QY 479 AAAAACTGATCTATTCGCAAGCATTTTGTATCATCTTTACTTTCCTGCTCAAG 538  
DB 2325 AAACTCTGTAATTAAGCAGCTGACTATAGGCTTGAATCAATTTGAACTTCCCTGACATT 2266  
QY 539 GAGAACTGAATGAG 553  
DB 2265 GAGTATGATTTGAG 2251

RESULT 4

US-08-545-528D-1  
Sequence 1, Application US/08545528D  
Patent No. 6537773  
GENERAL INFORMATION:  
APPLICANT: Fraser et al.  
TITLE OF INVENTION: Nucleotide Sequence of the Mycoplasma Genitalium Genome, Fragment  
Patent No. 6537773  
TITLE OF INVENTION: Thee of, and Uses Thereof  
FILE REFERENCE: PB193pl  
CURRENT APPLICATION NUMBER: US/08/545,528D  
CURRENT FILING DATE: 1995-10-19  
PRIOR APPLICATION NUMBER: US 08/488,018  
PRIOR FILING DATE: 1995-06-07  
PRIOR APPLICATION NUMBER: US 08/473,545  
PRIOR FILING DATE: 1995-06-07  
NUMBER OF SEQ ID NOS: 1  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1  
LENGTH: 580073  
TYPE: DNA  
ORGANISM: Mycoplasma genitalium  
US-08-545-528D-1

Query Match 6.2%; Score 34.4; DB 4; Length 580073;  
Best Local Similarity 48.9%; Pred. No. 17;  
Matches 92; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 366 TAAAGTCACTCATATGTTATACCAACGTTTCTGTTCCGAACTATACCATGAAATGC 425  
DB 281067 TGAATTAACATGATTAACCAACCAATTTTAAACCACTTGAATGCTTAAAGC 281126  
QY 426 AACATTTCTTAATATCTGAACCTGACGATGACCTGAAAAATTAGCTGAATAAC 485  
DB 281127 AACATTTCTGATATGATGATCACTAACAAATCCITTTAAGAAATATCATACGAAAAATC 281186  
QY 486 CTGATTTCTGCGCAAGCATTTTGTATCATCTTTACTTTCCTGCTGCAAGGAACT 545  
DB 281187 AAACGTAATTAAGCAATTTCTTTTAAACGATTTGAAAAAGATCATGTTCTGCTCAAAA 281246  
QY 546 GAATGAG 553  
DB 281247 AAAAGAG 281254

RESULT 5

US-09-134-001C-2533  
Sequence 2533, Application US/09134001C  
Patent No. 6380370  
GENERAL INFORMATION:  
APPLICANT: Lynn Doucette-Stamm et al  
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: GTC-007  
CURRENT APPLICATION NUMBER: US/09/134,001C  
CURRENT FILING DATE: 1998-08-13  
PRIOR APPLICATION NUMBER: US 60/064,964  
PRIOR FILING DATE: 1997-11-08  
PRIOR APPLICATION NUMBER: US 60/055,779  
PRIOR FILING DATE: 1997-08-14  
NUMBER OF SEQ ID NOS: 5674  
SEQ ID NO 2533  
LENGTH: 1500  
TYPE: DNA  
ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-2533

Query Match 6.1%; Score 33.8; DB 4; Length 1500;  
Best Local Similarity 54.4%; Pred. No. 2.6;  
Matches 68; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 102 CAGATACGAATACCTCTCTTCTATCTATCTATCTGAGTTCCTGTTGAAATG 161  
DB 146 CAGGCGCTATGCTTCTTCAATAGTATGATGATGATGATGATGATGATGATGATG 205  
QY 162 CAGCACTTTTCCGCAAGGCTGTAATAATAGCTTCAAGTCTGTCGATCTGCTT 221  
DB 206 CTGTAAGTATGAGGCACTGTTGAAAAATTCCTGTTCTGTTGTTGCGCATTTGCTTTA 265  
QY 222 GTCTT 226  
DB 266 GTT 270

RESULT 6

US-09-280-116-144  
Sequence 144, Application US/09280116A  
Patent No. 6331427  
GENERAL INFORMATION:  
APPLICANT: Robison, Keith E.  
TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs  
FILE REFERENCE: 5800-24, 035800/176965  
CURRENT APPLICATION NUMBER: US/09/280,116A  
CURRENT FILING DATE: 1999-03-26  
NUMBER OF SEQ ID NOS: 268  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 144  
LENGTH: 246  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: trypsin-like serine proteases  
US-09-280-116-144

Query Match 6.1%; Score 33.6; DB 4; Length 246;  
Best Local Similarity 52.9%; Pred. No. 1.5;  
Matches 72; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 286 CCAATGAGAACTCTGAGACATCTATATGACCCCAATTAATCTTCTCTGTTATG 345  
DB 38 CCAATGAGAGACAGAGGCTCTCCCAATGACCCAGGATATGATGATGATGATGATG 97  
QY 346 ATCCATTCCTTGTAGCTTTAAGTCACTCATATGATATACCAACAGCTTCTGTT 405  
DB 98 CTGCTCCACAGTTCATTTGCTGTTATCTCTACTGAGTGAAGGCTGACACTGATC 157  
QY 406 CCACTATCAACATGA 421  
DB 158 CCGGTTGCACTATGA 173

RESULT 7

US-09-497-855A-37  
Sequence 37, Application US/09497855A  
Patent No. 6605432  
GENERAL INFORMATION:

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 19, 2004, 20:57:37 ; Search time 2538 Seconds

(without alignments)  
9443.931 Million cell updates/sec

Title: US-09-531-113-5981

Perfect score: 553  
Sequence: 1 gtacgtatcttcacagacatt.....ccaagagagactgatgag 553

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 3470272 seqs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenBank  
1: gb\_ba:\*  
2: gb\_ba:\*  
3: gb\_in:\*  
4: gb\_in:\*  
5: gb\_ov:\*  
6: gb\_ov:\*  
7: gb\_ov:\*  
8: gb\_ov:\*  
9: gb\_ov:\*  
10: gb\_ov:\*  
11: gb\_ov:\*  
12: gb\_ov:\*  
13: gb\_ov:\*  
14: gb\_ov:\*  
15: gb\_ov:\*  
16: gb\_ov:\*  
17: gb\_ov:\*  
18: gb\_ov:\*  
19: gb\_ov:\*  
20: gb\_ov:\*  
21: gb\_ov:\*  
22: gb\_ov:\*  
23: gb\_ov:\*  
24: gb\_ov:\*  
25: gb\_ov:\*  
26: gb\_ov:\*  
27: gb\_ov:\*  
28: gb\_ov:\*  
29: gb\_ov:\*  
30: gb\_ov:\*  
31: gb\_ov:\*  
32: gb\_ov:\*  
33: gb\_ov:\*  
34: gb\_ov:\*  
35: gb\_ov:\*  
36: gb\_ov:\*  
37: gb\_ov:\*  
38: gb\_ov:\*  
39: gb\_ov:\*  
40: gb\_ov:\*  
41: gb\_ov:\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	133.8	24.2	1004	8	AK066056
C 2	125.8	22.7	1088	8	AK065722
C 3	123.2	22.3	591	8	AY098968
C 4	123.2	22.3	789	8	AF462824
C 5	98.6	17.8	277	6	AR245689
C 6	58	10.5	105858	2	AP004324
C 7	56.4	10.2	41087	8	AB024029
C 8	53.6	9.7	105967	8	AP005647
C 9	53.6	9.7	133387	2	AP004061
C 10	53.2	9.6	7218	6	166494
C 11	45.2	8.2	163277	2	AC074259
C 12	45.2	8.2	252394	3	AE014833
C 13	45	8.1	126038	2	AC016068
C 14	43.2	7.8	564	11	BV089971
C 15	42	7.6	218313	2	AC112682
C 16	42	7.6	307150	1	CNSPAX01
C 17	42	7.6	349980	6	AX041106
C 18	41.8	7.6	408	6	AX300942
C 19	41.6	7.5	192541	2	AC079165
C 20	41.6	7.5	198181	2	AC035248
C 21	41.6	7.5	215063	2	AC107763
C 22	41.2	7.5	258880	2	AC095430
C 23	41	7.4	246724	2	AC095185
C 24	41	7.4	267434	2	AC128998
C 25	40.8	7.4	224956	2	AC019008
C 26	40.8	7.4	231316	2	AC115138
C 27	40.8	7.4	251096	10	AL586255
C 28	40.2	7.3	1141	6	AX083744
C 29	40.2	7.3	16876	2	EX255937
C 30	40.2	7.3	181925	2	AC020710
C 31	40.2	7.3	197864	9	AC021956
C 32	40	7.2	53385	2	AC136994
C 33	40	7.2	125733	10	AC117250
C 34	40	7.2	140176	9	AC022188
C 35	40	7.2	152779	9	AC018943
C 36	40	7.2	187842	9	AC114778
C 37	40	7.2	217619	2	AC135511
C 38	39.8	7.2	164603	2	AC119466
C 39	39.8	7.2	223996	2	AC094504
C 40	39.6	7.2	142565	10	AL772205
C 41	39.6	7.2	188713	2	AC128580
C 42	39.6	7.2	243075	2	AC097827
C 43	39.4	7.1	42942	7	AB045978
C 44	39.4	7.1	43095	6	BD245282
C 45	39.4	7.1	111873	9	AC073275

#### ALIGNMENTS

RESULT 1  
AK066056/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS

AK066056 1004 bp mRNA  
Oryza sativa (japonica cultivar-group) cDNA clone:J013052F01, full  
insert sequence.

AK066056  
AK066056.1 GI:32976074  
FUT-CDNA; CAB trapper.  
Oryza sativa (japonica cultivar-group)  
Oryza sativa (japonica cultivar-group)  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliopsida; Liliopsida; Poales; Poaceae;  
Ehrhartoideae; Oryzeae; Oryza.

The Rice Full-length cDNA Consortium, National Institute of

Agrobiological Sciences Rice Full-Length cDNA Project Team:  
 Kikuchi, S., Sato, K., Nagata, T., Kawagashira, N., Doi, K.,  
 Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,  
 Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,  
 Ohtsuki, K., Shishiki, T., Foundation of Advancement of International  
 Science Genome Sequencing & Analysis Group, Otsu, Y., Murakami, K.,  
 Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,  
 Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,  
 Nakikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Nishikura, J.,  
 Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,  
 Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN,  
 Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,  
 Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,  
 Kagawa, I., Kondo, S., Komono, H., Miyazaki, A., Otsu, N., Ota, Y.,  
 Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,  
 Yoshino, M., and Hayashizaki, Y.  
 Collection, mapping, and annotation of over 28,000 cDNA clones from  
 japonica rice  
 Science 301 (5631), 376-379 (2003)

TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 REFERENCE  
 AUTHORS

2 (bases 1 to 1004)  
 12869764  
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,  
 Fujimura, T., Fukuda, S., Hara, A., Hashidume, W.,  
 Hayashizaki, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hirooka, T.,  
 Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K.,  
 Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,  
 Kikuchi, S., Kishikawa, Hirozane, T., Kishimoto, N., Kobayashi, M.,  
 Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Komono, H., Kouda, M.,  
 Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,  
 Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,  
 Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,  
 Namiki, T., Nakikawa, R., Nishikura, J., Nishi, K., Nomura, K.,  
 Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H.,  
 Otsu, N., Ota, Y., Otsu, Y., Ryu, R., Saito, H., Sakai, C., Sakai, K.,  
 Sakazume, N., Sano, H., Sasaki, D., Sato, K., Sato, K., Shibata, K.,  
 Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,  
 Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,  
 Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A.,  
 Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W.,  
 Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S., and  
 Yoshimura, A.  
 Direct Submission  
 Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of  
 Agrobiological Sciences, Department of Molecular Genetics, Head of  
 Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki  
 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp,  
 Tel:81-29-858-7007, Fax:81-29-858-7007)  
 This clone is one of the 28K full-length cDNA clones from japonica  
 rice.  
 URL: <http://cdna01.dna.affrc.go.jp/cDNA/>  
 NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Sato, K.,  
 Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,  
 Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T.,  
 Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., and  
 Yamamoto, M.  
 FAIS Genome Sequencing & Analysis Group: Otsu, Y., Iida, Y.,  
 Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,  
 Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J.,  
 Mizuno, K., Nakikawa, R., Nishikura, J., Oka, M., Ryu, R., Sugano, S.,  
 Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,  
 Yoshimura, A., Matsubara, K., and Murakami, K.  
 Genome Exploration Research Group in Riken Genomic Sciences Center  
 and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,  
 Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hara, A.,  
 Hara, A., Hashidume, W., Hayashizaki, K., Hayatsu, N., Hiramoto, K.,  
 Hirooka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,  
 Itoh, M., Kagawa, I., Kanagawa, S., Kato, H., Kawai, J.,  
 Kikuchi, S., Kishikawa, Hirozane, T., Kojima, Y., Kondo, S., Komono, H.,  
 Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Miura, J.,  
 Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Otsu, N.,  
 Ota, Y., Saito, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,

TITLE  
 JOURNAL  
 MEDLINE  
 PUBMED  
 REFERENCE  
 AUTHORS

Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,  
 Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,  
 Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K.,  
 Yasunishi, A., and Hayashizaki, Y.  
 Location/Qualifiers  
 1. 1004  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /mol\_type="mRNA"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:39947"  
 /clone="J013052P01"

ORIGIN  
 Query Match 24.2%; Score 133.8; DB 8; Length 1004;  
 Best Local Similarity 65.7%; Pred. No. 1,4e-21;  
 Matches 195; Conservative 0; Mismatches 102; Indels 0; Gaps 0;  
 QY 132 CTATCTATCTCTGAGTCTCTGTCGATTCGACCAACTTTCCCGAAGAGCTGTAAAAA 191  
 |||||  
 Db 642 CTATCTATCTCTGAGTCTCTGTCGATTCGACCAACTTTCCCGAAGAGCTGTAAAAA 583  
 QY 192 TACCTTCAAGTCTGTCGATTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 251  
 582 CGACTTCAATCTCTGTCGATTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 523  
 QY 252 CAGCTTCATCTCCAGTCGACGATCTGTCGATCTGTCGATCTGTCGATCTGTCGATCTGTC 311  
 522 AATCTTATTTGACCTCTGATCTGTCGATCTGTCGATCTGTCGATCTGTCGATCTGTCGAT 463  
 QY 312 ATGACCCCAATTAATCT 371  
 462 ATAGCTTCACT 403  
 QY 372 CAATCTATCTGTAACCAACAGCTTTCTTTTCCGAATCTGTCGATCTGTCGATCTGTCGAT 428  
 Db 402 CAGCTTATTAATTAATCT 346  
 RESULT 2  
 AK065722/c  
 LOCUS  
 DEFINITION  
 Oryza sativa (japonica cultivar-group) cDNA clone:J013038d22, full  
 insert sequence.  
 ACCESSION  
 AK065722  
 VERSION  
 AK065722.1 GI:32975740  
 KEYWORDS  
 FLI cDNA; CAP trapper.  
 SOURCE  
 Oryza sativa (japonica cultivar-group)  
 ORGANISM  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.  
 REFERENCE  
 1 The Rice Full-length cDNA Consortium, National Institute of  
 Agrobiological Sciences Rice Full-length cDNA Project Team:  
 Kikuchi, S., Sato, K., Nagata, T., Kawagashira, N., Doi, K.,  
 Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,  
 Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,  
 Ohtsuki, K., Shishiki, T., Foundation of Advancement of International  
 Science Genome Sequencing & Analysis Group: Otsu, Y., Iida, Y.,  
 Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,  
 Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J.,  
 Mizuno, K., Nakikawa, R., Nishikura, J., Oka, M., Ryu, R., Sugano, S.,  
 Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,  
 Yoshimura, A., Matsubara, K., and Murakami, K.  
 Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN,  
 Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,  
 Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,  
 Kagawa, I., Kondo, S., Komono, H., Miyazaki, A., Otsu, N., Ota, Y.,  
 Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,  
 Yoshino, M., and Hayashizaki, Y.  
 Collection, mapping, and annotation of over 28,000 cDNA clones from  
 japonica rice  
 Science 301 (5631), 376-379 (2003)

PUBMED  
REFERENCE  
AUTHORS

12869764  
2 (bases 1 to 1088)  
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Doi,K., Fujimura,T., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayashizaki,Y., Hayatsu,N., Hiramoto,K., Hiroaka,T., Hori,F., Hotta,I., Iida,J., Iida,Y., Ikeda,R., Imanura,K., Imotani,K., Ishibiki,J., Ishii,Y., Ishikawa,M., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawagashira,N., Kawai,J., Kawamata,M., Kikuchi,S., Kishikawa-Hirozane,T., Kishimoto,N., Kobayashi,M., Kodama,T., Kojima,K., Kojima,Y., Kondo,S., Kono,H., Kouda,M., Koyas,S., Kurihara,C., Kurosaki,T., Kusumegi,T., Li,C., Lu,M., Maeda,H., Matsubara,K., Matsuyama,T., Mura,J., Miyazaki,A., Mizuno,K., Murakami,K., Murata,M., Nagata,T., Nakamura,M., Namiki,T., Narikawa,R., Nishikura,J., Nishi,K., Nomura,K., Numasaki,R., Ohneda,E., Ohno,M., Ohtsuki,K., Oka,M., Ooka,H., Oseko,N., Ota,Y., Otsomo,Y., Ryu,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Satoh,K., Shibata,K., Shinagawa,A., Shiraki,T., Shishiki,T., Sogabe,Y., Sugano,S., Sugiyama,A., Suzuki,K., Suzuki,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Tsunoda,Y., Ueda,M., Waki,K., Xie,Q., Yahagi,M., Yamada,H., Yamamoto,M., Yasunishi,A., Yazaki,J., Yokomizo,S. and Yoshimura,A.  
Direct Submission  
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)  
This clone is one of the 28k full-length cDNA clones from japonica rice.  
URL: http://cdna01.dna.affrc.go.jp/cDNA/  
NIAS Rice Full-length cDNA Project Team: Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K., Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I., Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C., Ohtsuki,K., Shishiki,T. and Yamamoto,M.  
FAIS Genome Sequencing & Analysis Group: Otsomo,Y., Iida,Y., Fujimura,T., Ikeda,R., Ishibiki,J., Kawamata,M., Kobayashi,M., Kodama,T., Kurosaki,T., Kusumegi,T., Lu,M., Maeda,H., Mura,J., Mizuno,K., Narikawa,R., Nishikura,J., Oka,M., Sano,H., Sugano,S., Yoshimura,A., Suzuki,Y., Tsunoda,Y., Ueda,M., Xie,Q., Yokomizo,S., Yoshimura,A., Matsubara,K. and Murakami,K.  
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P., Fukuda,S., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiroaka,T., Hori,F., Iida,J., Iida,Y., Imanura,K., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kanagawa,S., Katoh,H., Kawai,J., Kishikawa-Hirozane,T., Kojima,Y., Kondo,S., Kono,H., Kouda,M., Nakamura,M., Nishikura,C., Matsuyama,T., Miyazaki,A., Murata,M., Ota,Y., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Sato,K., Shibata,K., Shingawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagami-Takeda,Y., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T., Waki,K., Yasunishi,A. and Hayashizaki,Y.  
Location/Qualifiers  
1. 1088  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="mRNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/clone="J013038J22"

FEATURES  
source

## ORIGIN

Query Match 22.7%; Score 125.8; DB 8; Length 1088;  
Best Local Similarity 64.0%; P-Ed. No. 1,1e-19;  
Matches 190; Conservative 0; Mismatches 107; Indels 0; Gaps 0;

OY 133 TATCATCTGAGTTCCTGTTGCAATGACGACCACTTTCCGGAACGCTGTAAAT 192  
DB 837 TACCATCTTGAGTCTCTTTCGAAATGCGCAGCTTTCTCCGATAGGTGAAGGAG 778

OY 193 AGCTTCAAGTCGTTGGCGATCTGTGTCTTGTCTTGATGCAAGATACCTTGCTCACTC 252  
DB 777 GACCTGAGAGCTCTTGCAATCTGCTCTTGTCGTGATGAGGAGCCCTTGATCGGTA 718  
OY 253 AGCTCATCTTCACCTGGAAGTCATCTAGTTCACCAATGGAATCTGGGACATCTTA 312  
DB 717 AACCTTACTTCTTACTCTCAAGTCCGACGTCACCAAGAAATCTCGGAATGTCCAGA 658  
OY 313 TGACCCCAATTAACCTTCTCTCTCTGATGATTCATTCATCCCTTGACTTTAAGTC 372  
DB 657 TGCCCTTGACCTTCATCTTCTTCTTCTTATTAATCACTTACCCTTAATTTCAAGCT 598  
OY 373 AACTCATATGATTAACAACAAGCTTCTGTTCCGAACTATCAACATGAATCAACA 429  
DB 597 AGTTATATGATATACCTACTCTCTTCTTATTAACAAGGTGACTAGAAGGATCA 541

RESULT 3  
AY098968/c  
LOCUS  
DEFINITION  
Arabidopsis thaliana AT5G58110/K2119\_90 mRNA, complete cds.  
ACCESSION  
AY098968  
VERSION  
AY098968.1 GI:20466112  
KEYWORDS  
SOURCE  
ORGANISM  
FLU cDNA.  
Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosid II; Brassicales; Brassicaceae; Arabidopsis.  
1 (bases 1 to 591)  
Kim,C.J., Chen,H., Cheuk,R., Meyers,M.C., Shim,P., Banh,J., Bowser,L., Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y., Ishida,J., Jones,T., Kamliya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
Arabidopsis ORF clones  
Unpublished  
2 (bases 1 to 591)  
Kim,C.J., Chen,H., Cheuk,R., Meyers,M.C., Shim,P., Banh,J., Bowser,L., Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y., Ishida,J., Jones,T., Kamliya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.  
Direct Submission  
Submitted (22-APR-2002) Salk Institute Genomic Analysis Laboratory (SIGAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA

## COMMENT

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPF cDNAs (RAPF cDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamliya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGSC (SSP) Consortium members carried out the sequencing and annotation of the RAPF cDNAs: Kim,C.J., Chen,H., Cheuk,R., Meyers,M.C., Shim,P., Banh,J., Bowser,L., Chang,E., Dale,J.M., Goldsmith,A.D., Jones,T., Karlin-Neumann,G., Lam,B., Lee,J.M., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Davis,R.W., Theologis,A., and Ecker,J.R.  
Kim,C.J. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk) contributed equally to this work as PI.

## FEATURES

source

Location/Qualifiers  
 1..591  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /chromosome="5"  
 /clone="U21471"  
 /ecotype="Columbia"  
 /note="This clone is in pUNI 51"

## CDS

1..591  
 /note="unknown protein"  
 /codon\_start=1  
 /product="AT5G58110/K21119\_90"  
 /protein\_id="AAU1978.1"  
 /db\_xref="GI:20466113"  
 /translation="MERERETKRGSSRYVWRATSDAAPPIPKLSNNDVSLNTPASIGSLMNRAGTWEEKSLTKWATDRLEKELGVSGLQFSGKAEIIDVNCVGDATLVTNNKRVGVGYEELSLKVEGWSFEENMKVKSLTEIPFSFGEIDLDLVDVLSLSDKELSQOKRIKIDMKQFLBPIRLKGLFPEGLKDR"

## ORIGIN

Query Match 22.3%; Score 123.2; DB 8; Length 591;  
 Best Local Similarity 63.5%; Pred. No. 5e-19;  
 Matches 186; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 133 TATCTATCTGAGTTCCTGTTGCAATGACGAACTTTCCGACAGGCTGTAAAT 192  
 |||||  
 Db 590 TATCTATCTGAGTTCCTGTTGCAAGCGCGAGTTGAGCGGATAGCTCCAGAAC 531  
 QY 193 AGCTTCAAGTGTGGCATCTGTGCTTCTTGATGACGAATATCCCTTGCTTCACTC 252  
 |||||  
 Db 530 TGCTTCAATCCAACTTAATCCGCTGCTGAGCTGCGGAAAGCTCCTGTCTGCTG 471  
 QY 253 AGTCTCCTTCACCTGCAAGTCAATCTAGTCCACCAATGAGACTCGGACATCTATA 312  
 |||||  
 Db 470 AGCTTCAATCCAACTTAATCCGCTGCTGAGCTGCGGAAAGCTCCTGTGCTG 411  
 QY 313 TGACCCCAATAAATCTTCTCTCTGATGATCCATCCCTTTGACTTTTAAGTTC 372  
 |||||  
 Db 410 CTCCTTCACCTTCTTCAATTCCTTCAATGACCACTCCCTCAACTTTAAGGAA 351  
 QY 373 AACTCATATGATTAACCAACAGCTTTCTTTCGAACTATCAACATTAATGCAAC 428  
 |||||  
 Db 350 AGCTGTAAGTATAGCCAACTCTTTCTTCTGCGAACTGACTTAAGAAAGCATC 295

## RESULT 4

AF462824/c 789 bp mRNA linear PLN 08-JAN-2002

LOCUS AF462824  
 DEFINITION Arabidopsis thaliana AT5G58110/K21119\_90 mRNA, complete cds.  
 ACCESSION AF462824  
 VERSION AF462824.1 GI:18087564  
 KEYWORDS FLI CDNA.

## SOURCE

ORGANISM Arabidopsis thaliana (thale cress)

## REFERENCE

AUTHORS Cheuk,R., Chen,H., Kim,C.J., Meyers,M.C., Banh,J., Bowser,L., Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinzaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.

TITLE Arabidopsis cDNA clones  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 789)

## AUTHORS

Cheuk,R., Chen,H., Kim,C.J., Meyers,M.C., Banh,J., Bowser,L., Carninci,P., Chang,E., Dale,J.M., Goldsmith,A.D., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J.,

## TITLE

JOURNAL

## COMMENT

Lam,B., Lee,J.M., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Satou,M., Seki,M., Southwick,A., Tang,C.C., Toriumi,M., Wu,H.C., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Shinzaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.  
 Direct Submission  
 Submitted (27-DEC-2001) Salk Institute Genomic Analysis Laboratory (SIGAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA  
 RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RFL cDNAs (RFL cDNA: "RIKEN Arabidopsis Full-Length cDNA"): Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinzaki,K.

The Salk, Stanford, PSEC (SSP) Consortium members carried out the sequencing and annotation of the RFL cDNAs: Cheuk,R., Chen,H., Kim,C.J., Meyers,M.C., Shinn,P., Banh,J., Bowser,L., Dale,J.M., Goldsmith,A.D., Jiang,P.X., Jones,T., Karlin-Neumann,G., Lam,B., Lee,J.M., Lin,J., Liu,S.X., Miranda,M., Nguyen,M., Onodera,C.S., Palm,C.J., Pham,P.K., Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Yamada,K., Yamamura,Y., Yu,G., Yu,S., Davis,R.W., Theologis,A., and Ecker,J.R.

Cheuk,R. (SSP/Salk) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinzaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/Salk) contributed equally to this work as PIs.

## FEATURES

source

Location/Qualifiers  
 1..789  
 /organism="Arabidopsis thaliana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3702"  
 /chromosome="5"  
 /clone="RAF109-09-K03 (R21471)"  
 /ecotype="Columbia"  
 1..33  
 34..624  
 /note="unknown protein"  
 /codon\_start=1  
 /product="AT5G58110/K21119\_90"  
 /protein\_id="AAU58914.1"  
 /db\_xref="GI:18087565"  
 /translation="MERERETKRGSSRYVWRATSDAAPPIPKLSNNDVSLNTPASIGSLMNRAGTWEEKSLTKWATDRLEKELGVSGLQFSGKAEIIDVNCVGDATLVTNNKRVGVGYEELSLKVEGWSFEENMKVKSLTEIPFSFGEIDLDLVDVLSLSDKELSQOKRIKIDMKQFLBPIRLKGLFPEGLKDR"

## CDS

QY 133 TATCTATCTGAGTTCCTGTTGCAATGACGAACTTTCCGACAGGCTGTAAAT 192  
 |||||  
 Db 623 TATCTATCTGAGTTCCTGTTGCAAGCGCGAGTTGAGCGGATAGCTCCAGAAC 564  
 QY 193 AGCTTCAAGTGTGGCATCTGTGCTTCTTGATGACGAATATCCCTTGCTTCACTC 252  
 |||||  
 Db 563 TGCTTCAATCCAACTTAATCCGCTGCTGAGCTGCGGAAAGCTCCTGTCTGCTG 504  
 QY 253 AGTCTCCTTCACCTGCAAGTCAATCTAGTCCACCAATGAGACTCGGACATCTATA 312  
 |||||  
 Db 503 AGCTTCAATCCAACTTAATCCGCTGCTGAGCTGCGGAAAGCTCCTGTGCTG 444  
 QY 313 TGACCCCAATAAATCTTCTCTCTGATGATCCATCCCTTTGACTTTTAAGTTC 372  
 |||||  
 Db 443 CTCCTTCACCTTCTTCAATTCCTTCAATGACCACTCCCTCAACTTTAAGGAA 384  
 QY 373 AACTCATATGATTAACCAACAGCTTTCTTTCGAACTATCAACATTAATGCAAC 428  
 |||||  
 Db 383 AGCTGTAAGTATAGCCAACTCTTTCTTCTGCGAACTGACTTAAGAAAGCATC 328

## ORIGIN

3'UTR

Query Match 22.3%; Score 123.2; DB 8; Length 789;  
 Best Local Similarity 63.5%; Pred. No. 4.9e-19;  
 Matches 186; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

RESULT 5  
AR245689/c  
LOCUS AR245689 277 bp DNA linear PAT 20-DEC-2002  
DEFINITION Sequence 1048 from patent US 6476212.  
ACCESSION AR245689  
VERSION AR245689.1 GI:27293563  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 277)  
AUTHORS Laigudi,R.V., Ito,L.Y. and Sherman,B.K.  
TITLE Polynucleotides and polypeptides derived from corn ear  
JOURNAL Patent: US 6476212-A 1048 05-NOV-2002;  
FEATURES  
Location/Qualifiers  
1..277  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 17.8%; Score 98.6; DB 6; Length 277;  
Best Local Similarity 61.5%; Pred. No. 4e-13;  
Matches 158; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 167 ACTTTCCTCCGAGAGGCTGTAAATAGCTCAAGTCGTCGATCTGTGCTT 226  
DB 257 ATTTTCTCTGAGATCGGCGCCAAAGAGCTTGAATCTTGAATCTTATCAT 198  
QY 227 GATGCAAGATATCCCTTGTCTACTCACTCACTCCAGCTGCAATCATATGTTAC 286  
DB 197 CAGACGAGAGGCTTGTGATCGCGCAACTTATCTTGAATCTTGAATCATCGATCGC 138  
QY 287 CAATGAGAACTCTGGGACATCTATGACCCCAATTAATCTTCTCTCTGTATGA 346  
DB 137 CGAATGAAACTCGGGAATGTAGATGCCCTTGTGATCTTCTTCTTGTATTA 78  
QY 347 TCCATTCCTCTTGTACCTTAAAGTCACTATGATATAACAAGCTTCTTGTTC 406  
DB 77 ACCATTACCGCTAAATCTCAAGCTTATGTTATATCTTCTCTTGTTC 18  
QY 407 GAACATATCACCATGAT 423  
DB 17 GGACATGACCTGAAT 1

RESULT 6  
AP004324/c  
LOCUS AP004324 105858 bp DNA linear HTG 21-MAR-2002  
DEFINITION Oryza sativa (japonica cultivar-group) chromosome 6 clone  
OI1215\_E11, \*\*\* SEQUENCING IN PROGRESS \*\*\*  
ACCESSION AP004324  
VERSION AP004324.1 GI:16930109  
KEYWORDS HTG; HTGS\_PHASE2.  
SOURCE Oryza sativa (japonica cultivar-group)  
ORGANISM Oryza sativa (japonica cultivar-group)  
REFERENCE Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Eriatoidae; Oryzae; Oryza.  
1  
REFERENCE Sasaki,T., Matsumoto,T. and Yamamoto,K.  
AUTHORS Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, BAC  
TITLE clone:OI1215\_E11  
JOURNAL Published Only in Database (2001)  
REFERENCE 2 (bases 1 to 105858)  
AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.  
TITLE Direct Submission  
JOURNAL Submitted (14-NOV-2001) Takuji Sasaki, National Institute of  
Agricultural Sciences, Rice Genome Research Program, Kannondai  
2-1-2, Tsukuba, Ibaraki 305-8602, Japan  
(E-mail: sasaki@nias.affrc.go.jp, URL: http://rnp.dna.affrc.go.jp/  
Tel: 81-298-38-7441, Fax: 81-298-38-7468)

COMMENT  
The nucleotide sequence of this BAC clone was generated by  
combining Monsanto and RGP-Japan sequencing data.  
NOTE: It currently consists of 1 contigs. Gaps between the contigs  
are represented as runs of N. The order of the pieces is believed  
to be correct as given, however the sizes of the gaps between them  
are based on estimates that have provided by the submitter. This  
sequence will be replaced by the finished sequence as soon as it is  
available and the accession number will be preserved.  
\* NOTE: This is a 'working draft' sequence.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.

FEATURES  
source  
Location/Qualifiers  
1..105858  
/organism="Oryza sativa (japonica cultivar-group)"  
/mol\_type="genomic DNA"  
/cultivar="Nipponbare"  
/db\_xref="taxon:39947"  
/chromosome="6"  
/clone="OI1215\_E11"

ORIGIN  
Query Match 10.5%; Score 58; DB 2; Length 105858;  
Best Local Similarity 62.3%; Pred. No. 0.0013;  
Matches 91; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 132 CTATCTATCTGTAGTCTCTGTTGCAATTCAGCACTTCCGAGAGGCTGTAAAA 191  
DB 68554 CTATCTATCTGTAGTCTCTGTTGCAATTCAGCACTTCCGAGAGGCTGTAAAA 68495  
QY 192 TAGCTTCAAGTGTGGGATCTGTGTCTTGTCTTGAAGATATCCCTTGTACT 251  
DB 68494 CGACTTCAATCTTGTGATATTTGTGCTTGTGTGCGAGAGGCTTGTATGTGCT 68435  
QY 252 CAGTCTACTTCCAGCTGCAATCAT 277  
DB 68434 AAAATTTATTTGCACTGAAAGAAAT 68409

RESULT 7  
AB024029/c  
LOCUS AB024029 41087 bp DNA linear PLN 27-DEC-2000  
DEFINITION Arabidopsis thaliana genomic DNA, chromosome 5, TAC clone:K21119.  
ACCESSION AB024029  
VERSION AB024029.1 GI:4519188  
KEYWORDS  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
REFERENCE Arabidopsis thaliana  
AUTHORS Arabidopsis thaliana  
TITLE Structural analysis of Arabidopsis thaliana chromosome 5. X.  
JOURNAL Sequence features of the regions of 3,076,755 bp covered by sixty  
BAC and PAC clones.  
MEDLINE DNA Res. 7 (1), 31-63 (2000)  
PUBMED 10718197  
REFERENCE 2 (bases 1 to 41087)  
AUTHORS Nakamura,Y.  
TITLE Direct Submission  
JOURNAL Submitted (24-FEB-1999) Yasukazu Nakamura, Kazusa DNA Research  
Institute, Department of Plant Gene Research, 1532-3, Yana,  
Kisarazu, Chiba 292-0812, Japan (E-mail: ynakamura@kazusa.or.jp,  
Tel: 81-438-52-3935, Fax: 81-438-52-3934)  
Address for correspondence: kao@kazusa.or.jp  
For the latest information on annotation of this clone, please see  
http://www.kazusa.or.jp/kaos/cgi-bin/and\_graph.cgi?c=K21119  
Genes with similarity to proteins in the databases are described in  
'product' or 'note' qualifiers. Genes that have no significant  
protein similarity are described as 'unknown protein'.

The software programs used to predict genes include: Graal (Informatics Group, Oak Ridge National Laboratory, <http://compbio.ornl.gov/Graal-1.3/>), GENSCAN (Chris Burge, MIT, <http://CCR-081.mit.edu/GENSCAN.html>), NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of Denmark, <http://www.cbs.dtu.dk/services/NetGene2/>) and SplicePredictor (Volker Brendel, Stanford University, <http://gremjini.zool.iastate.edu/cgi-bin/np.cgi>). Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, Washington University School of Medicine, St. Louis, <http://genome.wustl.edu/eddy/tRNAscan-SE/>). This sequence may not be the entire insert of this clone. It may be shorter because we remove overlaps between neighboring submissions. The 5' clone is F2C19 and the 3' clone is MCK7.

## FEATURES

## Source

1. 41087  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/strain="Columbia"  
/db\_xref="taxon:3702"  
/chromosome="5"  
/clone="K21L19"  
/clone\_1lb="Mitau TAC"  
19. 864  
/note="gene\_id:F2C19.3  
ref:NP\_057491.1  
similar to unknown protein"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB10993.1"  
/db\_xref="GI:9759527"  
/translation="MTLPRLRGGGGGGATGAERDCYLMNAYAKKPDVDPNEOR  
LSKMLCALSNELAEVCVLDLGNLKNKLVHALSKLPPOFVSIKAKLKMVNIK  
LTPVAGSDSSODTTSQFOCPVSGLEFNKKEFALRGCHVMAKALKEVSSCL  
VCAVDKSDKTYINGTEHEVDLRRMEKAKAKREKGVSKKNGAAVADTAK  
VAKQLEDGVNNGITVKKFKADKPVNAIKVYASLFTSKKSDPEFTYCSLS  
PLGRN"

## CDS

complement(join(1236..1268,1357..1464,1557..1623,  
1715..1791,1886..1963,2119..2182,2287..2375,2627..2698)  
/note="gb|AAD34459.1  
gene\_id:K21L19.2  
similar to unknown protein"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB10994.1"  
/db\_xref="GI:9759528"  
/translation="MIGVGMKQYNSVLDKPLSKGKQVSTAFAPLSEITVNOYNO  
VDNIALERRLEDAGYAVGSRVLELONREKRRRTRLGILSFVHSTVWKVLEKQV  
ADSLKGEHEHEDYIMSEKELVNRFTSLPKDMGTNCGAFVAGIVKGVLDNAGFPVAV  
VTAHFVPIBQOQRPRTTILIKFADEVILKREARLSQ"  
join(3080..3847,4376..4605,4728..4854,4933..5001,  
5132..5272,5425..5526,5640..7229,7426..7968)  
/note="gene\_id:K21L19.3  
unknown protein"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB10995.1"  
/db\_xref="GI:9759529"  
/translation="MEEDPEPDLYSGDVALQPPQPVVLPPEPLPHESTIDNLRSDO  
DYSERASIPRSNDNDAYKLSQDARTQIVDGGDDKMSFDIERPADSPTPTPG  
LFTGALPGLATDRGVSYVTRISQVGGGGGGGGGGGGDDSDSDSDITVIND  
SSRNWIGAGDRSRMGDNEDDEDDEDEPLVVAUTDPNQMEQMGEDGLQGIEG  
DGADGAGKSGSGGATGPPKAGYSGHGHVPSQKVPYRPGAAPFGAASVGGPS  
SGQVPPANIGPMAGRGGRGMDPLKMRNNAQAQGHQPGWSTAGRGDFTLLPST  
IEFVDDISFEKPMRYGVEMTDYFNGGLNEESWKQYCKQLDQHRIOQTMOGSLIRYE  
SGRTDGGVDPDLPEPLAATAGAGVDSNLTKPSQVDSKQKAPVANNRPPLPPEP  
IYVTSGERLPSITPRAPRMRLDAIIEHSHDEPSGNGTQALSSIPGENVPEET  
SYVNNKRPDTSASHSPADQEPKHLKQDDERISSTDSGQFRSSSPVGDGTSS  
SVYRVDVGEAGADAMGELKMSFTSPQAVDGDGKSTERSSESSKARSGSRD  
FOOEBVDIODKSSRPANNKQYDNNAPHSRKNQDQKEMERTTAAKSGRENSNDP  
MELDSTYISIASREDPDKREKEDVGAVMRKEDDPYSGGDEGSRKRDREDDGF  
ROGKRENRISKIDQVPSRKMDAGMNTIYEPDHIINKRKDEEVYRRSPKNE  
ISVGRSEMSRVRERDDRLRHQKRVYQHKIRDDPDRSGSLKQRDDITMQRDGNELR

## CDS

complement(join(1236..1268,1357..1464,1557..1623,  
1715..1791,1886..1963,2119..2182,2287..2375,2627..2698)  
/note="gb|AAD34459.1  
gene\_id:K21L19.2  
similar to unknown protein"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB10994.1"  
/db\_xref="GI:9759528"  
/translation="MIGVGMKQYNSVLDKPLSKGKQVSTAFAPLSEITVNOYNO  
VDNIALERRLEDAGYAVGSRVLELONREKRRRTRLGILSFVHSTVWKVLEKQV  
ADSLKGEHEHEDYIMSEKELVNRFTSLPKDMGTNCGAFVAGIVKGVLDNAGFPVAV  
VTAHFVPIBQOQRPRTTILIKFADEVILKREARLSQ"  
join(3080..3847,4376..4605,4728..4854,4933..5001,  
5132..5272,5425..5526,5640..7229,7426..7968)  
/note="gene\_id:K21L19.3  
unknown protein"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB10995.1"  
/db\_xref="GI:9759529"  
/translation="MEEDPEPDLYSGDVALQPPQPVVLPPEPLPHESTIDNLRSDO  
DYSERASIPRSNDNDAYKLSQDARTQIVDGGDDKMSFDIERPADSPTPTPG  
LFTGALPGLATDRGVSYVTRISQVGGGGGGGGGGGGDDSDSDSDITVIND  
SSRNWIGAGDRSRMGDNEDDEDDEDEPLVVAUTDPNQMEQMGEDGLQGIEG  
DGADGAGKSGSGGATGPPKAGYSGHGHVPSQKVPYRPGAAPFGAASVGGPS  
SGQVPPANIGPMAGRGGRGMDPLKMRNNAQAQGHQPGWSTAGRGDFTLLPST  
IEFVDDISFEKPMRYGVEMTDYFNGGLNEESWKQYCKQLDQHRIOQTMOGSLIRYE  
SGRTDGGVDPDLPEPLAATAGAGVDSNLTKPSQVDSKQKAPVANNRPPLPPEP  
IYVTSGERLPSITPRAPRMRLDAIIEHSHDEPSGNGTQALSSIPGENVPEET  
SYVNNKRPDTSASHSPADQEPKHLKQDDERISSTDSGQFRSSSPVGDGTSS  
SVYRVDVGEAGADAMGELKMSFTSPQAVDGDGKSTERSSESSKARSGSRD  
FOOEBVDIODKSSRPANNKQYDNNAPHSRKNQDQKEMERTTAAKSGRENSNDP  
MELDSTYISIASREDPDKREKEDVGAVMRKEDDPYSGGDEGSRKRDREDDGF  
ROGKRENRISKIDQVPSRKMDAGMNTIYEPDHIINKRKDEEVYRRSPKNE  
ISVGRSEMSRVRERDDRLRHQKRVYQHKIRDDPDRSGSLKQRDDITMQRDGNELR

## CDS

complement(join(1236..1268,1357..1464,1557..1623,  
1715..1791,1886..1963,2119..2182,2287..2375,2627..2698)  
/note="gb|AAD34459.1  
gene\_id:K21L19.2  
similar to unknown protein"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB10994.1"  
/db\_xref="GI:9759528"  
/translation="MIGVGMKQYNSVLDKPLSKGKQVSTAFAPLSEITVNOYNO  
VDNIALERRLEDAGYAVGSRVLELONREKRRRTRLGILSFVHSTVWKVLEKQV  
ADSLKGEHEHEDYIMSEKELVNRFTSLPKDMGTNCGAFVAGIVKGVLDNAGFPVAV  
VTAHFVPIBQOQRPRTTILIKFADEVILKREARLSQ"  
join(3080..3847,4376..4605,4728..4854,4933..5001,  
5132..5272,5425..5526,5640..7229,7426..7968)  
/note="gene\_id:K21L19.3  
unknown protein"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB10995.1"  
/db\_xref="GI:9759529"  
/translation="MEEDPEPDLYSGDVALQPPQPVVLPPEPLPHESTIDNLRSDO  
DYSERASIPRSNDNDAYKLSQDARTQIVDGGDDKMSFDIERPADSPTPTPG  
LFTGALPGLATDRGVSYVTRISQVGGGGGGGGGGGGDDSDSDSDITVIND  
SSRNWIGAGDRSRMGDNEDDEDDEDEPLVVAUTDPNQMEQMGEDGLQGIEG  
DGADGAGKSGSGGATGPPKAGYSGHGHVPSQKVPYRPGAAPFGAASVGGPS  
SGQVPPANIGPMAGRGGRGMDPLKMRNNAQAQGHQPGWSTAGRGDFTLLPST  
IEFVDDISFEKPMRYGVEMTDYFNGGLNEESWKQYCKQLDQHRIOQTMOGSLIRYE  
SGRTDGGVDPDLPEPLAATAGAGVDSNLTKPSQVDSKQKAPVANNRPPLPPEP  
IYVTSGERLPSITPRAPRMRLDAIIEHSHDEPSGNGTQALSSIPGENVPEET  
SYVNNKRPDTSASHSPADQEPKHLKQDDERISSTDSGQFRSSSPVGDGTSS  
SVYRVDVGEAGADAMGELKMSFTSPQAVDGDGKSTERSSESSKARSGSRD  
FOOEBVDIODKSSRPANNKQYDNNAPHSRKNQDQKEMERTTAAKSGRENSNDP  
MELDSTYISIASREDPDKREKEDVGAVMRKEDDPYSGGDEGSRKRDREDDGF  
ROGKRENRISKIDQVPSRKMDAGMNTIYEPDHIINKRKDEEVYRRSPKNE  
ISVGRSEMSRVRERDDRLRHQKRVYQHKIRDDPDRSGSLKQRDDITMQRDGNELR

## CDS

complement(join(8713..9011,9095..9578,9651..9790,  
9879..10199,10276..10774,10865..10974,11057..11353,  
11490..11601)  
/note="gene\_id:K21L19.4  
pir||104792  
strong similarity to unknown protein"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB10996.1"  
/db\_xref="GI:9759530"

## CDS

complement(join(12492..12684,12967..13064,13154..13258,13475..13540,  
13752..13838,14293..14343)  
/note="gb|AAD00112.1; vesicle trafficking protein  
SNARE-like  
gene\_id:K21L19.5"  
/codon\_start=1  
/evidence=not experimental  
/product="ATGp1"  
/protein\_id="BAB10997.1"  
/db\_xref="GI:9759531"

## CDS

complement(join(14727..14912,14997..15371)  
/note="gene\_id:K21L19.6"  
/codon\_start=1  
/evidence=not experimental  
/product="outer membrane lipoprotein-like"  
/protein\_id="BAB10998.1"  
/db\_xref="GI:9759532"  
/translation="MTKEKMEVVGKLANVERVYMGWYEIASPFRFQPNKGVDTNAY  
TINPQGTIVLNEVWNGKRGFIEGSAVYADKSDAKIKGVYVPEPLPIIPYGDY  
WLYIDPDYQHALIGQPSRSYTIMLSRTQMEETKQIVKRAVEGCIYISKLAKTPQ  
SDTPPSNTAPEDSKGVWFKSLFGK"  
complement(join(16000..16065,16143..17255,17352..17428,  
17558..17920,17997..18146,18236..18365))  
/note="contains similarity to two-component response  
regulator protein"  
gene\_id:K21L19.7"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB10999.1"  
/db\_xref="GI:9759533"

## CDS

complement(join(16000..16065,16143..17255,17352..17428,  
17558..17920,17997..18146,18236..18365))  
/note="contains similarity to two-component response  
regulator protein"  
gene\_id:K21L19.7"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAB10999.1"  
/db\_xref="GI:9759533"  
/translation="MEFSGTEGRHDKPPVGMRYLAVDNPTCLRLLEILLRCKYH  
TMTMSRKALIMLRNSNMEDYIVDVEKPDGFTLLEIGLMDIPVTHSDYSVM  
KNIHGACDYIVKPGKSLQNTVHWVYKNTKSYAKLLPESSEDVPSASRRKRV  
NKGSDGDEDRDDEGESEGDGSGRKRVRVYSLQHPVAVQQLGDKVAV  
KILDLIMSIEGILTRNVASHLOKLYLKEIDEGQOQNTTPAFGRDVSYPQADL  
GRLDLMASQIPSSGILSRSHLTKLOPKYISINIQANSSPFIQGHQNSNSANP  
FCTYSSTSPRIQVNTLPORSTSPLELOFPSPKSLIGFCKLGRALGAGSLDTCMP  
RSSSTSLPSAETNPLMIAQNTYQPLHLSADIOPICTBPSTNSAPNISFQGLSRP  
GSMWQGNLNTTRFPSSLPNLNLAFLPDQVTCAGNNIGDCTLVSANPGEQCDPQL

LGFMONVPLGGOKMEQONCTMLNPNNGNIEYPLPADNMVPRDNNSTRSKJDESLM  
 NPIDNSOEYVGATMTMDPEMNSGKPEENDNDVDVDDIMNEMKPEENNNGVPAVATRF  
 GFDSFPP  
 complement (join(18421, .18675,18945, .18992))  
 /note="gene.id:K21L19.8  
 unknown protein"  
 /codon\_start=1  
 /evidence=not\_experimental  
 /protein\_id="BAB11000.1"  
 /db\_xref="GI:9759534"  
 /translation="MNSNRSRIIRMPEDCQIRDSTNRKREIRGCDLSHLVLYGNMR  
 NRRKLIIMRIKQJQINRYITLISFGLPVLPHLSTIGLMSSSPPLVLRKQIRF"  
 complement (join(19771, .20875,21080, .21408))  
 /note="gene.id:K21L19.9"  
 /codon\_start=1  
 /evidence=not\_experimental  
 /product="glucanase; glucan endo-1,3-beta-glucosidase"  
 /protein\_id="BAB11001.1"  
 /db\_xref="GI:9759535"  
 /translation="MWSGSLVLLLAVALLCQASSIGANWGTQASHPLPDIIVAMLR  
 ENGLOKVLPAEYDITRALGSGIEVWVGIPNEMLATIASLKAERKVAKNVSTHI  
 STDNINIRYVAVGNPEPLSTYNGSYLSTFPALRNIGAIKAGIQNGVKTCPINAD

Query Match 10.2%; Score 56.4; DB 8; Length 41087;  
 Best Local Similarity 63.0%; Pred. No. 0.0034;  
 Matches 87; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 133 TATCATCTGAGTTCCTGTTGCAATGACGAACTTTCCCGAAGAGCTGTAAAT 192  
 DB 30693 TATCATCTTGAAGTCTGTTGCAAGAGCGCACTTTGAGCGGATATAGCTCCCAAGAC 30694  
 QY 193 AGCTTCAAGTCGTTGGCAGATCTGTCTTCTTGATGCAAGATATCCCTGCTTCACTC 252  
 DB 30633 TGCCTCATATCCAACTTAATCGCTTCGCTTCAGCTGCGAAGCTCTTGTCTGCTG 30574  
 QY 253 AGTCTCACTTCACCTGCG 270  
 DB 30573 AGCTTCACATCCACCTGCG 30556

RESULT 8  
 AP005647  
 LOCUS  
 DEFINITION Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 2,  
 BAC clone:OSUNBA002E05, complete sequence.  
 ACCESSION AP005647  
 VERSION AP005647.2 GI:31414614  
 KEYWORDS HTG;  
 SOURCE Oryza sativa (japonica cultivar-group)  
 ORGANISM Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Eukaryota; Viridiplantae; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE  
 AUTHORS Sasaki,T., Matsumoto,T. and Katayose,Y.  
 TITLE Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, BAC  
 clone:OSUNBA002E05  
 JOURNAL Published Only in Database (2002)  
 REFERENCE 2 (bases 1 to 105967)  
 AUTHORS Sasaki,T., Matsumoto,T. and Katayose,Y.  
 JOURNAL Direct Submission  
 Submitted (21-AUG-2002) Takuji Sasaki, National Institute of  
 Agrobiological Sciences, Rice Genome Research Program, Kamondai  
 2-1-2, Tsukuba, Ibaraki 305-8602, Japan  
 (E-mail:tsasaki@nias.affrc.go.jp, URL:http://rsgp.dna.affrc.go.jp/,  
 Tel:81-298-38-7441, Fax:81-298-38-7468)  
 COMMENT On Jun 4, 2003 this sequence version replaced gi:22415832.  
 The orientation of the sequence is from -21M13 to M13rev of the BAC  
 clone.

FEATURES  
 source  
 1..105967  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /mol\_type="genomic DNA"

/cultivar="Nipponbare"  
 /db\_xref="taxon:39947"  
 /chromosome="2"  
 /clone="OSUNBA002E05"

ORIGIN  
 Query Match 9.7%; Score 53.6; DB 8; Length 105967;  
 Best Local Similarity 61.4%; Pred. No. 0.015;  
 Matches 86; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 133 TATCATCTGAGTTCCTGTTGCAATGACGAACTTTCCCGAAGAGCTGTAAAT 192  
 DB 79127 TACCTATCTTGAAGCTTCTTCGATGCGCGAGTTCTCCCGATGAGTGAAGGAAG 79186  
 QY 193 AGCTTCAAGTCGTTGGCAGATCTGTCTTCTTGATGCAAGATATCCCTGCTTCACTC 252  
 DB 79187 GACTTGAAGCTCTTGAATATGCTCTTGTGTCGATGAGAGCCCTTGTCATCGGTA 79246  
 QY 253 AGTCTCACTTCACCTGCAA 272  
 DB 79247 AACCTACTTCACTGAAA 79266

RESULT 9  
 AP004061  
 LOCUS  
 DEFINITION Oryza sativa (japonica cultivar-group) chromosome 2 clone  
 OJ1407 E09, \*\*\* SEQUENCING IN PROGRESS \*\*\*.  
 ACCESSION AP004061  
 VERSION AP004061.1 GI:15208429  
 KEYWORDS HTG; PHASE2.  
 SOURCE Oryza sativa (japonica cultivar-group)  
 ORGANISM Oryza sativa (japonica cultivar-group)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Eukaryota; Viridiplantae; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE  
 AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.  
 TITLE Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, BAC  
 clone:OJ1407 E09  
 JOURNAL Published Only in Database (2001)  
 REFERENCE 2 (bases 1 to 133387)  
 AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.  
 JOURNAL Direct Submission  
 Submitted (15-AUG-2001) Takuji Sasaki, National Institute of  
 Agrobiological Resources, Rice Genome Research Program, Kamondai  
 2-1-2, Tsukuba, Ibaraki 305-8602, Japan  
 (E-mail:tsasaki@nias.affrc.go.jp, URL:http://rsgp.dna.affrc.go.jp/,  
 Tel:81-298-38-7441, Fax:81-298-38-7468)  
 COMMENT The nucleotide sequence of this BAC clone was generated by  
 combining Monsanto and RGP-Japan sequencing data.  
 NOTE: It currently consists of 1 contigs. Gaps between the contigs  
 are represented as runs of N. The order of the pieces is believed  
 to be correct as given, however the sizes of the gaps between them  
 are based on estimates that have provided by the submitter. This  
 sequence will be replaced by the finished sequence as soon as it is  
 available and the accession number will be preserved.  
 \* NOTE: This is a 'working draft' sequence.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.

FEATURES  
 source  
 1..133387  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /mol\_type="genomic DNA"  
 /cultivar="Nipponbare"  
 /db\_xref="taxon:39947"  
 /chromosome="2"  
 /clone="OJ1407\_E09"

ORIGIN  
 Query Match 9.7%; Score 53.6; DB 2; Length 133387;  
 Best Local Similarity 61.4%; Pred. No. 0.015;



Matches	86;	Conservative	0;	Mismatches	54;	Indels	0;	Gaps	0;
QY	133	TATCATCTGAGATTCCCTGTTGGAATTCAGACAACTTTTCCCGAACAGCGCTATAAAAT	192						
Db	132967	TACCATCTCTTGAAGCTCTTCTTCGATGCGCAGATTCTCCCGGATAGGTGAAGAAG	133026						
QY	193	AGCTTCAGTGGTGGGAGATCTGTGCTCTGTGATGGAAGTATCCCTGGTTGACATC	252						
Db	133027	GACTTGAAGTCTCTTCAAAATCTGCCCTCTTGTGTCGGATGCGAGGCCCTTGTATGCTA	133086						
QY	253	AGTCTCACTTCACCTGCAA	272						
Db	133087	AACTTACTTCTTACCTGAAA	133106						

LOCUS	166494	7218 bp	DNA	linear	PAT 28-DEC-1997
DEFINITION	Sequence 14 from patent US 5670367.				
ACCESSION	166494				
VERSION	166494.1	GI:2724471			
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	1 (bases 1 to 7218)				
AUTHORS	Dorner, F., Scheifflinger, F. and Falkner, F. Gunter.				
TITLE	Recombinant fowlpox virus				
JOURNAL	Patent: US 5670367-A 14 23-SEP-1997;				
FEATURES	Location/Qualifiers				
source	1..7218				

Query Match 9.6%; Score 53.2; DB 6; length 7218;  
 Best Local Similarity 3.4%; Pred. No. 0.02;  
 Matches 13; Conservative 219; Mismatches 152; Indels 0; Gaps 0;

LOCUS	AC074259	163277 bp	DNA linear	HTG 17-JUL-2001
DEFINITION	Trypanosoma brucei chromosome VI clone PC193-4M18, *** SEQUENCING IN PROGRESS ***.			

ACCESSION	AC074259
VERSION	AC074259.13 GI:14787212
KEYWORDS	HTG; HTGS_PHASE2.
SOURCE	Trypanosoma brucei
ORGANISM	Trypanosoma brucei

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS

1 (bases 1 to 163777)  
El-Sayed,N.M., Ghedin,E., Song,J., Larkin,C., Wanless,D., Jones,K.,  
Peterson,S., Hou,L., Zhao,H., Mason,T., Miltchevcher,J., Pai,G., Van  
Aken,S., Uetreback,T., Kralak,H.G., Gerard,C., Leech,V., Ullu,E.,  
Melville,S., White,C., Adams,M.D., Donelson,J.F. and Fraser,C.M.  
Trypanosoma brucei GUTat10.1 RPi193-4M18 BAC genomic sequence  
Unpublished  
2 (bases 1 to 163777)  
El-Sayed,N.M., Khalak,H. and Adams,M.D.

```

FEATURES
    source
        Location/Qualifiers
            1..163277
                /organism="Trypanosoma brucei"
                /mol_type="genomic DNA"
                /isolate="GUTat10.1"
                /db_xref="taxon:5691"
                /chromosome="VI"
                /clone="RPC193-4M18"

```

RESULT	12
AE014833	
LOCUS	
DEFINITION	252394 bp DNA linear INV 07-OCT-2002
ACCESSION	AE014833
VERSION	AE014833 AE014185
KEYWORDS	complete sequence.
SOURCE	AE014833.1 GI:23495098
ORGANISM	.
REFERENCE	Plasmodium falciparum 3D7 Bukaryova; Alviolata; Apicomplexa; Haemosporida; Plasmodium. 1 (bases 1 to 252394)
AUTHORS	Gardner,M.J., Hall,N., Fung,E., White,O., Berriman,M., Hyman,R.W., Carlton,J.M., Paine,A., Nelson,K.E., Bowman,S., Paulsen,I.T., James,K., Eisen,J.A., Rutherford,K., Salzberg,S.L., Craig,A., Kyes,S., Chan,M.-S., Nene,V., Shallom,S.J., Suh,B., Peterson,J.,

Angiuoli, S., Perlea, M., Allen, J., Selengut, J., Haft, D.,  
Mather, M.W., Vaidya, A.B., Martin, D.M.A., Fairlamb, A.H.,  
Fraunholz, M.J., Roos, D.S., Ralph, S.A., McFadden, G.I.,  
Cummins, L.M., Subramanian, G.M., Mungall, C., Venter, J.C.,  
Carucci, D.J., Hoffman, S.L., Newbold, C., Davis, R.W., Fraser, C.M. and  
Barrell, B.  
TITLE  
Genome sequence of the human malaria parasite *Plasmodium falciparum*  
JOURNAL  
Nature 415 (6906), 498-511 (2002)  
PUBMED  
12368864  
REFERENCE  
2 (bases 1 to 252394)  
AUTHORS  
Gardner, M.J.  
TITLE  
Direct Submision  
JOURNAL  
Submitted (13-SEP-2002) The Institute for Genomic Research, 9712  
Medical Center Dr., Rockville, MD 20850, USA  
FEATURES  
Location/Qualifiers  
source  
1. 252394

/organism="Plasmodium falciparum 3D7"  
/mol\_type="genomic DNA"  
/isolate="3D7"  
/db\_xref="taxon:36329"  
/chromosome="10"  
complement(79..213)  
/note="AT rich"  
/rpt\_type=tandem  
253..301  
repeat\_region  
/rpt\_type=tandem  
complement(253..301)  
/rpt\_type=tandem  
/rpt\_unit="(TA)n"  
318..374  
repeat\_region  
/rpt\_type=tandem  
324..374  
repeat\_region  
/rpt\_type=tandem  
/rpt\_unit="(TAA)n"  
complement(433..482)  
/rpt\_type=tandem  
/rpt\_unit="(TA)n"  
577..605  
repeat\_region  
/rpt\_type=tandem  
complement(577..605)  
/rpt\_type=tandem  
/rpt\_unit="(A)n"  
complement(661..682)  
repeat\_region  
/rpt\_type=tandem  
/rpt\_unit="(TA)n"  
complement(792..942)  
/rpt\_type=tandem  
/rpt\_unit="(TA)n"  
897..1030  
repeat\_region  
/rpt\_type=tandem  
/rpt\_unit="(TATAA)n"  
complement(1038..1173)  
/note="AT rich"  
/rpt\_type=tandem  
1100..1130  
repeat\_region  
/rpt\_type=tandem  
complement(1100..1126)  
/rpt\_type=tandem  
/rpt\_unit="(TA)n"  
1133..1172  
repeat\_region  
/rpt\_type=tandem  
1175..1207  
repeat\_region  
/rpt\_type=tandem  
complement(1175..1207)  
/rpt\_type=tandem  
/rpt\_unit="(A)n"  
complement(1218..1281)  
/note="AT rich"  
/rpt\_type=tandem  
1365..1391  
repeat\_region  
/rpt\_type=tandem  
complement(1365..1391)  
/rpt\_type=tandem

repeat\_region  
complement(1401..1493)  
/note="AT rich"  
/rpt\_type=tandem  
1505..1526  
repeat\_region  
/rpt\_type=tandem  
/rpt\_unit="(A)n"  
complement(1602..1685)  
/rpt\_type=tandem  
/rpt\_unit="(TA)n"  
complement(1726..1825)  
/note="AT rich"  
/rpt\_type=tandem  
complement(1963..2007)  
/note="AT rich"  
/rpt\_type=tandem  
complement(2049..2157)  
/rpt\_type=tandem  
/rpt\_unit="(TA)n"  
2064..2131  
repeat\_region  
/rpt\_type=tandem  
complement(2194..2390)  
/rpt\_type=tandem  
/rpt\_unit="(TA)n"  
2257..2332  
repeat\_region  
/rpt\_type=tandem  
complement(2405..2457)  
/note="AT rich"  
/rpt\_type=tandem  
complement(2481..2516)  
/note="AT rich"  
/rpt\_type=tandem  
2532..2615  
repeat\_region  
/rpt\_type=tandem  
/rpt\_unit="(A)n"  
complement(2623..2657)  
/note="AT rich"  
/rpt\_type=tandem  
2666..2699  
repeat\_region  
/rpt\_type=tandem  
/rpt\_unit="(TAA)n"  
complement(2740..2807)  
/note="AT rich"  
/rpt\_type=tandem  
2855..3047  
repeat\_region  
/rpt\_type=tandem  
complement(2856..3047)  
/rpt\_type=tandem  
/rpt\_unit="(TA)n"  
2893..3047  
repeat\_region  
/rpt\_type=tandem  
3100..3137  
repeat\_region  
/rpt\_type=tandem  
complement(3100..3137)  
/rpt\_type=tandem  
/rpt\_unit="(TA)n"  
complement(3145..3170)  
/note="AT rich"  
/rpt\_type=tandem  
complement(3196..3240)  
/note="AT rich"  
/rpt\_type=tandem  
3283..3319  
repeat\_region  
/rpt\_type=tandem  
complement(3287..3319)  
/rpt\_type=tandem  
/rpt\_unit="(TA)n"  
complement(3367..3434)  
/note="AT rich"  
/rpt\_type=tandem  
complement(3513..3656)  
/rpt\_type=tandem  
/rpt\_unit="(TA)n"

```

repeat_region complement (3659. .3794)
/rpt_type=tandem
/rpt_unit="(TA)n"
repeat_region complement (3809. .3838)
/note="AT_rich"
/rpt_type=tandem
/rpt_unit="(A)n"
repeat_region complement (3861. .3861)
/rpt_type=tandem
/rpt_unit="(A)n"
repeat_region complement (3866. .4026)
/note="AT_rich"
/rpt_type=tandem
/rpt_unit="(A)n"
repeat_region complement (4089. .4110)
/note="AT_rich"
/rpt_type=tandem
/rpt_unit="(TA)n"
repeat_region complement (4251. .4414)
/note="AT_rich"
/rpt_type=tandem
/rpt_unit="(TA)n"
repeat_region complement (4262. .4295)
/rpt_type=tandem
/rpt_unit="(TA)n"
repeat_region complement (4361. .4384)
/rpt_type=tandem
/rpt_unit="(TA)n"
repeat_region complement (4424. .4483)
/rpt_type=tandem
/rpt_unit="(TA)n"
repeat_region complement (4511. .4553)
/rpt_type=tandem
/rpt_unit="(TA)n"
repeat_region complement (4602. .4628)
/note="AT_rich"
/rpt_type=tandem
/rpt_unit="(TA)n"
repeat_region complement (4663. .4725)
/note="AT_rich"

```

Query Match 8.2%; Score 45.2; DB 3; Length 252394;  
 Best Local Similarity 52.7%; Pred. No. 1.4;  
 Matches 98; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

```

QY 58 TACAAGTACGAGTAAATTAAGCAAAACGACGCTTGTGCAGATACGAAATCTTC 117
DB 234162 TACAAATATGATAGACATGATATTAATTTCTTTTAAAGATTATTCATCTTT 234221
QY 118 CTCTTCTATCTATCTATCTATCTGAGTTCCTGTCGAAATTCAGCAACTTTCCCGA 177
DB 234222 ATATATATATATATATATATATATATATATATATATATATATATATATATAT 234281
QY 178 ACAGGCTGTAATAATAGCTTCAAGTGGTGGCATCTGTGTCTTGTTCGACGAAGA 237
DB 234282 AAAAACCCTTAAGATTAATATTCGATACATTATATGTGATTAACCTTATGCAAT 234341
QY 238 TCCCTT 243
DB 234342 TATTTT 234347

```

RESULT 13  
 ACO16068/c 126038 bp DNA linear HTG 13-JUL-2000  
 LOCUS Homo sapiens chromosome 4 clone RP11-431B5 map 4, LOW-PASS SEQUENCE  
 DEFINITION SAMPLING.  
 ACCESSION ACO16068  
 VERSION ACO16068.3 GI:9126605  
 KEYWORDS HTG; HTGS PHASE0.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

# REFERENCE

1 (bases 1 to 126038)  
 Birren, B., Linton, L., Nusbaum, C., and Lander, E.  
 Homo sapiens chromosome 4, clone RP11-431B5  
 Unpublished  
 2 (bases 1 to 126038)

# JOURNAL

Baldwin, J., Barna, N., Beckert, R., Boguslavsky, L., Bonkhalter, B.,  
 Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A.,  
 Cooke, P., Devereaux, K., Dewar, K., Domingo, M., Donelan, J., Doyle, M.,  
 Ferreira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D.,  
 Galagan, J., Gardyna, S., Grant, G., Hages, B., Heath, A., Horton, L.,  
 Howland, J., Johnson, R., Jones, C., Kamm, L., Karas, A., Klein, T.,  
 Leoczky, J., Lien, C., Locke, K., McDonald, P., Marquis, N.,  
 McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrum, J.,  
 Morrow, J., Naylor, J., Norman, C., O'Connor, T., O'Donnell, P.,  
 Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,  
 Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,  
 Testa, S., Tittell, A., Vasilev, H., Vo, A., Wheeler, J., Wu, X.,  
 Wyman, D., Ye, M., Zimmer, A. and Zody, M.  
 Direct Submission  
 Submitted (20-NOV-1999) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Jul 13, 2000 this sequence version replaced gi:5715901.  
 All repeats were identified using RepeatMasker:  
 Smith, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: MIBR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu

Project Information  
 Center project name: L4278  
 Center clone name: 431\_B5

\* NOTE: This record contains 130 individual  
 \* sequencing reads that have not been assembled into  
 \* contigs. Runs of N are used to separate the reads  
 \* and the order in which they appear is completely  
 \* arbitrary. Low-pass sequence sampling is useful for  
 \* identifying clones that may be gene-rich and allows  
 \* overlap relationships among clones to be deduced.  
 \* However, it should not be assumed that this clone  
 \* will be sequenced to completion. In the event that  
 \* the record is updated, the accession number will  
 \* be preserved.

```

1
890 889: contig of 889 bp in length
990 889: gap of 100 bp
1895 1894: contig of 905 bp in length
1895 1894: gap of 100 bp
1995 2876: contig of 882 bp in length
2877 2977 3868: gap of 100 bp
3869 3969 4868: contig of 892 bp in length
4869 4969 5882: contig of 900 bp in length
4969 5882: gap of 100 bp
5883 5882: contig of 914 bp in length
5883 5882: gap of 100 bp
5883 6879: contig of 897 bp in length
6879 6979: gap of 100 bp
6979 7908: contig of 929 bp in length
7908 8008: gap of 100 bp
8009 8908: contig of 900 bp in length
8908 9008: gap of 100 bp
9009 9870: contig of 862 bp in length
9871 9971 10879: contig of 909 bp in length
9971 10879: gap of 100 bp
10880 10979: contig of 884 bp in length
10980 11863: gap of 100 bp
11864 12838: contig of 875 bp in length
12838 12938: gap of 100 bp

```

\* 12939 13829: contig of 891 bp in length  
\* 12930 13929: gap of 100 bp  
\* 13930 14854: contig of 925 bp in length  
\* 14855 14954: gap of 100 bp  
\* 14955 15863: contig of 909 bp in length  
\* 15864 15963: gap of 100 bp  
\* 15964 16950: contig of 887 bp in length  
\* 16951 17871: contig of 921 bp in length  
\* 17872 18858: contig of 887 bp in length  
\* 18859 18959: gap of 100 bp  
\* 18959 19790: contig of 732 bp in length  
\* 19791 20687: contig of 897 bp in length  
\* 20688 20787: gap of 100 bp  
\* 20788 21671: contig of 884 bp in length  
\* 21672 21771: gap of 100 bp  
\* 21772 22673: contig of 902 bp in length  
\* 22674 22773: gap of 100 bp  
\* 22774 23665: contig of 892 bp in length  
\* 23666 23765: gap of 100 bp  
\* 23766 24662: contig of 897 bp in length  
\* 24663 24762: gap of 100 bp  
\* 24763 25690: contig of 928 bp in length  
\* 25691 25791: gap of 100 bp  
\* 25791 26697: contig of 906 bp in length  
\* 26697 26796: gap of 100 bp  
\* 26797 27710: contig of 914 bp in length  
\* 27711 27811: gap of 100 bp  
\* 27811 28693: contig of 883 bp in length  
\* 28694 28793: gap of 100 bp  
\* 28794 29700: contig of 907 bp in length  
\* 29701 29800: gap of 100 bp  
\* 29801 30699: contig of 899 bp in length  
\* 30700 30799: gap of 100 bp  
\* 30799 31699: contig of 900 bp in length  
\* 31700 31799: gap of 100 bp  
\* 31799 32718: contig of 919 bp in length  
\* 32719 32818: gap of 100 bp  
\* 32819 33744: contig of 926 bp in length  
\* 33745 33844: gap of 100 bp  
\* 33845 34734: contig of 890 bp in length  
\* 34735 34834: gap of 100 bp  
\* 34835 35771: contig of 937 bp in length  
\* 35772 35871: gap of 100 bp  
\* 35872 36789: contig of 918 bp in length  
\* 36790 36889: gap of 100 bp  
\* 36890 37797: contig of 908 bp in length  
\* 37798 37897: gap of 100 bp  
\* 37898 38819: contig of 922 bp in length  
\* 38820 38919: gap of 100 bp  
\* 38920 39854: contig of 935 bp in length  
\* 39855 39954: gap of 100 bp  
\* 39955 40860: contig of 906 bp in length  
\* 40861 40960: gap of 100 bp  
\* 40961 41848: contig of 888 bp in length  
\* 41849 41948: gap of 100 bp  
\* 41949 42837: contig of 889 bp in length  
\* 42838 42937: gap of 100 bp  
\* 42938 43842: contig of 905 bp in length  
\* 43843 43942: gap of 100 bp  
\* 43943 44850: contig of 908 bp in length  
\* 44851 44950: gap of 100 bp  
\* 44951 45883: contig of 933 bp in length  
\* 45884 45983: gap of 100 bp  
\* 45984 46879: contig of 896 bp in length  
\* 46880 46979: gap of 100 bp  
\* 46980 47858: contig of 879 bp in length  
\* 47859 47958: gap of 100 bp  
\* 47959 48839: contig of 881 bp in length  
\* 48840 48939: gap of 100 bp  
\* 48940 49851: contig of 912 bp in length

\* 49852 49951: gap of 100 bp  
\* 49952 50886: contig of 935 bp in length  
\* 50887 50986: gap of 100 bp  
\* 50987 51916: contig of 930 bp in length  
\* 51917 52017: gap of 100 bp  
\* 52017 52913: contig of 836 bp in length  
\* 52913 53013: gap of 100 bp  
\* 53013 53911: contig of 899 bp in length  
\* 53912 54011: gap of 100 bp  
\* 54012 54890: contig of 879 bp in length  
\* 54891 54990: gap of 100 bp  
\* 54991 55881: contig of 891 bp in length  
\* 55882 55981: gap of 100 bp  
\* 55982 56858: contig of 877 bp in length  
\* 56859 56958: gap of 100 bp  
\* 56959 57860: contig of 902 bp in length  
\* 57861 57960: gap of 100 bp  
\* 57961 58859: contig of 839 bp in length  
\* 58860 58959: gap of 100 bp  
\* 58960 59855: contig of 896 bp in length  
\* 59856 59955: gap of 100 bp  
\* 59956 60872: contig of 917 bp in length  
\* 60873 60972: gap of 100 bp  
\* 60972 61865: contig of 893 bp in length  
\* 61866 61965: gap of 100 bp  
\* 61966 62867: contig of 902 bp in length  
\* 62867 62967: gap of 100 bp  
\* 62968 63861: contig of 894 bp in length  
\* 63862 63961: gap of 100 bp  
\* 63962 64858: contig of 887 bp in length  
\* 64859 64958: gap of 100 bp  
\* 64959 65851: contig of 893 bp in length  
\* 65852 65951: gap of 100 bp  
\* 65952 66849: contig of 898 bp in length  
\* 66850 66949: gap of 100 bp  
\* 66950 67838: contig of 885 bp in length  
\* 67839 67938: gap of 100 bp  
\* 67939 68873: contig of 935 bp in length  
\* 68874 68973: gap of 100 bp  
\* 68974 69883: contig of 910 bp in length  
\* 69884 69983: gap of 100 bp  
\* 69984 70890: contig of 907 bp in length  
\* 70891 70990: gap of 100 bp  
\* 70991 71930: contig of 940 bp in length  
\* 71931 72030: gap of 100 bp

Query Match 8.1%; Score 45; DB 2; Length 126038;  
Best Local Similarity 50.7%; Pred. No. 1.7;  
Matches 108; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 282 TTACCAATGAGACTCTGGACATCTATATGACCCCAATAAATCTTCTCCTTG 341  
Db 95436 TTCCCGAGTAAGAACCTTTACCAAAAATTTATCCAGAAATATCCAGCCAGG 95377  
QY 342 TATGATCATTCCTCCCTTTGACTTTAAGTCATCTATGATATACCAACAGCTTCTT 401  
Db 95376 TATTCAGATATTCATTTAGGTGTCTGTTCTTAAGTTCCTTGAACAGGTTCTT 95317  
QY 402 GTTCGCAATGATGCAATGATGCAATTTCTTAATAATTCGACCTGACGATACCT 461  
Db 95316 GTCCCTTCTTTGATTTATTCATGAAATTTGATTAATTTGAATTAAGGTCAGGTTCCCT 95257  
QY 462 TGAATAATTTAGCTGAAAAAACCTGATTCTA 494  
Db 95256 TCAACTAATTTTATTTGAATAATATGTAATTTTA 95224

RESULT 14  
BV089971/c 564 bp DNA linear STS 15-OCT-2003  
LOCUS BV089971  
DEFINITION RPAMSE0001867 Roche Palo Alto Mus musculus STS genomic, sequence  
tagged site.  
ACCESSION BV089971

VERSION BV089971.1 GI:37667450  
 STS.  
 KEYWORDS Mus musculus (house mouse)  
 SOURCE Mus musculus  
 ORGANISM Mus musculus  
 REFERENCE 1 (bases 1 to 564)  
 AUTHORS Usuka, J., Iiao, G., Cheng, J., Nguyen, A., Bach, C., Puech, A., McPherson, J. D., Fournier, D. and Peltz, G.  
 TITLE Mus musculus SNPs  
 JOURNAL Unpublished (2003)  
 COMMENT  
 Contact: Jonathan Usuka  
 Roche Palo Alto Genetics and Genomics Department  
 Roche Palo Alto  
 3431 Hillview Ave, Mailstop 83-1, Palo Alto, CA 94024, USA  
 Tel: 6508555807  
 Email: jonathan.usuka@roche.com  
 Primer A: No primer submitted  
 Primer B: No primer submitted.  
 Location/Qualifiers  
 1..564  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /map="5-8454-8031-CAA01214661.1.1.43014"  
 /clone\_id="Roche Palo Alto"  
 /note="SNPs developed from assay sequences derived from 15 different strains of mice (as of October 1, 2003). Those strains include A/J, A/Heu, ~129/Sv, AKR/J, B10.D2-H2/Osnu, BALB/cBy, BALB/cJ, C3H/HeJ, C57BL/6J, CAST/Ei, DBA/2J, MRJ/MpJ, NZB/Blnd, NZW/Iac, SPRET/Ei."

ORIGIN  
 STS  
 Query Match 7.8%; Score 43.2; DB 11; Length 564;  
 Best Local Similarity 63.5%; Pval. No. 6.7;  
 Matches 66; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 40 AAGAGATTGCTAGTGAATCACTAGCTAATAAATTAACAATAAGTGGCTTTGT 99  
 Db 551 AAAGTATTGCTAATAAATCACTAATAAGTATCTCTGGTACCCAGTGGCCCTTTGT 492  
 QY 100 CACAGATGCAATATCTCTCTTCTATCTATCTATCTATCTCT 143  
 Db 491 CCGGTCCAAATTTTATCTATCTATCTATCTATCTATCTAT 448

RESULT 15  
 AC112682 218313 bp DNA linear HTG 05-DEC-2003  
 LOCUS Mus musculus chromosome 6 clone RP23-332K10 map 6, \*\*\* SEQUENCING  
 DEFINITION IN PROGRESS \*\*\*, 6 ordered pieces.  
 AC112682  
 AC112682.5 GI:38708110  
 HTG: HTGS\_PHRASE2; HTGS\_FULLTOP; HTGS\_ACTIVEFIN.  
 KEYWORDS Mus musculus (house mouse)  
 SOURCE Mus musculus  
 ORGANISM Mus musculus  
 REFERENCE 1 (bases 1 to 218313)  
 AUTHORS Birren, B., Nussbaum, C. and Lander, E.  
 TITLE Mus musculus chromosome 6, clone RP23-332K10  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 218313)  
 AUTHORS Birren, B., Linton, L., Nussbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearlano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,

TITLE JOURNAL  
 REFERENCE 3 (bases 1 to 218313)  
 AUTHORS Birren, B., Nussbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Corum, B., Dearlano, K., Diaz, J. S., Dodge, S., Doyle, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hagos, B., Haploplan, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Lander, E., Levine, R., Lindblad-Toh, K., Liu, X., Liu, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J., Menus, L., Mihova, T., Nguyen, C., Nicol, R., Norby, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieck, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Theodorou, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.  
 Direct Submission  
 Submitted (22-FEB-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
 3 (bases 1 to 218313)  
 Birren, B., Nussbaum, C., Lander, E., Abouelleil, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Corum, B., Dearlano, K., Diaz, J. S., Dodge, S., Doyle, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Graham, L., Grand-Pierre, N., Hagos, B., Haploplan, D., Hagos, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Lander, E., Levine, R., Lindblad-Toh, K., Liu, X., Liu, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J., Menus, L., Mihova, T., Nguyen, C., Nicol, R., Norby, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieck, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Theodorou, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.  
 Direct Submission  
 Submitted (05-DEC-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Dec 5, 2003 this sequence version replaced gi:27884887.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html  
 Direct Submission  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 Project Information  
 Center project name: L22646  
 Center clone name: 332\_K10

NOTE: This is a 'working draft' sequence. It currently consists of 6 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have provided by the submitter.  
 This sequence will be replaced  
 by the finished sequence as soon as it is available and  
 the accession number will be preserved.  
 1 67544: contig of 67544 bp in length  
 67545 67644: gap of 100 bp  
 67645 132358: contig of 64714 bp in length  
 132359 132458: gap of 100 bp  
 132459 157227: contig of 24769 bp in length  
 157228 157327: gap of 100 bp  
 157328 197939: contig of 40612 bp in length  
 197940 198039: gap of 100 bp  
 198040 203890: contig of 5851 bp in length  
 203891 203990: gap of 100 bp

FEATURES \* 203991 218313: contig of 14323 bp in length.  
Location/Qualifiers  
1..218313

/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="6"  
/map="6"  
/clone="RP23-332K10"  
/clone\_lib="RP23 Female Mouse BAC"

ORIGIN

Query Match 7.6%; Score 42; DB 2; Length 218313;  
Best Local Similarity 48.7%; Pred. No. 8.4; Indels 0; Gaps 0;  
Matches 114; Conservative 0; Mismatches 120;

QY 119 TCTTTCTATCTATCTATCTATCTGAGTTCCTGTTGCAATGCAACAACCTTTCCCGAA 178  
Db 88582 TCTATCTATCTATCTATCTATCTATCTATCTATCTGCTGCTGCTGCTATTTATTCAT 88641  
QY 179 CAGGCTGTAATAATAGCTTCAAGTGTGCGCATCTGTCTTGTCTTGATGCAAGATAT 238  
Db 88642 CTATCTATATATATCAACCTTCATATTTATCTATATTTATGACTGTTGAGCTAATAT 88701  
QY 239 CCTTGTCTCACTCACTCACTTCCAGCTGCAAGTCACTTATGTTCAACCAATGAGAACT 298  
Db 88702 ATATCCATCACTCAACACATATCTATCTAGCACTATCTATTTATCTATCTATCTACC 88761  
QY 299 CTGGGACATCTANATGACCCCAATAACTTCTCTCTCTGATGATGATCCATT 352  
Db 88762 TACCTACCTACCTATCTGTTACTATATATCACTATATTTATATATACACCTTT 88815

Search completed: March 19, 2004, 22:27:39  
Job time : 2539 secs



QY 228 ATGCAAGATATCCCTGCTTCATCTCAGTCTTCATCCAGCTGCAAGTCACTAGTACC 287  
Db 18057 CTTGATGATGCTGCTTTTCTTTTAACTTCCTTCAGGTTGTGTTAAGATG 17998  
QY 288 AATGAGACCTCGGACATATATAGACCCCAATAACTTCTTCTCTCTATGAT 347  
Db 17997 TCTTCTGATATCTGGATATAAGATATCTCATATCTTCTGCTCTCTATTT 17938  
QY 348 CCATTCCTTCGACTTTAA 368  
Db 17937 AATTTTAATTCATTTGAAA 17917  
RESULT 14  
US-09-790-988-1/c  
Sequence 1, Application US/09790988  
Patent No. 6632935  
GENERAL INFORMATION:  
APPLICANT: SHIGENOBU, SHUJI  
APPLICANT: MATSUNABE, HIDEKI  
APPLICANT: HATTORI, MASAHIRA  
APPLICANT: SAKAKI, YOSHIYUKI  
TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS  
FILE REFERENCE: 081356/0159  
CURRENT APPLICATION NUMBER: US/09/790,988  
CURRENT FILING DATE: 2001-02-23  
PRIOR APPLICATION NUMBER: JP2000-107160  
PRIOR FILING DATE: 2000-04-07  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 640681  
TYPE: DNA  
ORGANISM: Buchnera sp.  
US-09-790-988-1  
Query Match 6.0%; Score 33; DB 4; Length 640681;  
Best Local Similarity 49.2%; Pred. No. 45;  
Matches 87; Conservative 0; Mismatches 90; Indels 0; Gaps 0;  
QY 225 TTGATGCAAGATATCCCTGCTTCATCTCAGTCTTCATCCAGCTGCAAGTCACTAGTTC 284  
Db 507116 TTTCTGATGATTTTTCCTGCTGTTTCATGATCCATCTTACCAACATATATGTTATTT 507057  
QY 285 ACCAATGAGACTCTGGGACATCTATATGACCCCAATAACTTCTCTCTGAT 344  
Db 507056 TGACATATAACATCTACATAGATTGCTGCTGTAATAATTCACCTAATTTACTAA 506997  
QY 345 GATCCATTCCTCTTACTTTTAAAGTCAACTCATATGATATACCAACAGCTTTCTT 401  
Db 506996 GTTCTATGCTCTTTTCATGTAATAATATGTTCAATTTTACCAACATATTTT 506940  
RESULT 15  
US-08-916-421B-1/c  
Sequence 1, Application US/08916421B  
Patent No. 6503729  
GENERAL INFORMATION:  
APPLICANT: Buit et al.  
TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanoco-  
Patent No. 6503729  
TITLE OF INVENTION: jannaschi  
FILE REFERENCE: PB275  
CURRENT APPLICATION NUMBER: US/08/916,421B  
CURRENT FILING DATE: 1997-08-22  
PRIOR APPLICATION NUMBER: US 60/024,428  
PRIOR FILING DATE: 1996-08-22  
NUMBER OF SEQ ID NOS: 3  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1  
LENGTH: 1664976  
TYPE: DNA  
ORGANISM: Methanococcus jannaschi

FEATURE:  
NAME/KEY: misc feature  
LOCATION: (28222)..(28222)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (28257)..(28258)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (84773)..(84773)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (84808)..(84808)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (84812)..(84812)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (98120)..(98120)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (98159)..(98159)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (98239)..(98239)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (98266)..(98266)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (98343)..(98343)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (103998)..(103998)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (148948)..(148948)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (163385)..(163385)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (191989)..(191989)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (191995)..(191995)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (231980)..(231980)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (234187)..(234187)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (234220)..(234220)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (234814)..(234814)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (309398)..(309398)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (309418)..(309418)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (312837)..(312837)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (312993)..(312993)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (319226)..(319226)  
OTHER INFORMATION: n equals a, t, c, or g



```
NAME/KEY: misc_feature
LOCATION: (559167)..(559167)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (559241)..(559241)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (600992)..(600992)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (622708)..(622708)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (657081)..(657081)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (657203)..(657203)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (674435)..(674435)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (682442)..(682442)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (713652)..(713652)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (741684)..(741684)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (779455)..(779455)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (779676)..(779676)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (855539)..(855539)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (871619)..(871619)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1084830)..(1084830)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1096846)..(1096846)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1119881)..(1119881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1130881)..(1130881)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1310988)..(1310988)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1313224)..(1313224)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1345473)..(1345473)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1349491)..(1349491)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1470091)..(1470091)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1569020)..(1569020)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
```

```
LOCATION: (1602912)..(1602912)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1603734)..(1603734)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1637998)..(1637998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc_feature
LOCATION: (1664854)..(1664854)
OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1
```

Query Match 6.0%; Score 33; DB 4; Length 1664976;  
Best Local Similarity 46.7%; Pred. NO. 60;  
Matches 105; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

```
QY 207 GCGCATCTGTGTCTTCTTGTATGCAAGATATCCCTTGCTTCACTCACTGCTTCCAC 266
DB 1436207 GCACATGTGAATGGGATTTAAATTATATAGCTTTTAAACAATATCTTGCCCTTAT 1436148
QY 267 CTGCAAGTCATCTAGTTTACCAAAATGAGACTCTGGACATCTATATGACCCCAATAA 326
DB 1436147 CTTTCCATCTCTTCTTCTTCCCTTCTTATATATCTTCACTCTTGAACATCAATAT 1436088
QY 327 CTTCTCTCTCTTGTATGATTCATTCCTTCTTGAATTTAAGTCAATCATATATATA 386
DB 1436087 CCTACTCTTTAACAATGATTTCTTGCTTTTACATGAAACAATCTTACAGGTTTTC 1436028
QY 387 ACCAAGCTTTCTTGTTCGCAACTATACCATGAATGAACATT 431
DB 1436027 AGACGTATTTAATGCTTCAATTTTCAATTAATTCCTTT 1435983
```

Search completed: March 19, 2004, 21:39:31  
Job time : 76 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 19, 2004, 20:09:56 ; Search time 2134 Seconds

(without alignments)  
7738.419 Million cell updates/sec

Title: US-09-531-113-5981

Perfect score: 553  
Sequence: 1 gtaggtatttcagacatc.....ccaaggagaaatgagatggag 553

Scoring table: IDENTITY NUC  
Gapop 10\_0, Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*  
1: em\_estdb:\*  
2: em\_estdb:\*  
3: em\_estdb:\*  
4: em\_estdb:\*  
5: em\_estdb:\*  
6: em\_estdb:\*  
7: em\_estdb:\*  
8: em\_estdb:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_est3:\*  
12: gb\_est4:\*  
13: gb\_est5:\*  
14: gb\_est6:\*  
15: em\_estdb:\*  
16: em\_estdb:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_hum:\*  
20: em\_gss\_hum:\*  
21: em\_gss\_hum:\*  
22: em\_gss\_hum:\*  
23: em\_gss\_hum:\*  
24: em\_gss\_hum:\*  
25: em\_gss\_hum:\*  
26: em\_gss\_hum:\*  
27: em\_gss\_hum:\*  
28: gb\_gss1:\*  
29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	371.6	67.2	645	10	AM349814 GM210005B
2	323.8	58.6	425	12	BM270700 sak16e09
3	305.4	55.2	513	13	BM785857 sak8e02
4	300.8	54.4	383	12	BG238159 sab62h12

Result No.	Score	Query Match	Length	ID	Description
5	279.2	50.5	432	10	AM620829 sj47a12.Y
6	272.8	49.3	375	13	B0964888 sat03h07
7	255.8	46.3	576	28	BH405680 GM_UMB001
8	228.2	41.3	619	10	BE943027 EST422606
9	228.2	41.3	637	14	CB891233 EST648202
10	228.2	41.3	680	12	BM812845 EST590938
11	228.2	41.3	730	12	BM812845 EST590938
12	228.2	41.3	730	12	BM812845 EST590938
13	227.8	40.5	532	28	AQ989424 GM_UMB001
14	223.8	40.5	712	13	B0548744 GM880015B
15	221.2	40.0	726	12	BG586953 EST488722
16	205.2	37.1	252	12	BG510324 sac76d01
17	205.2	37.1	605	12	BM308189 sak42g12
18	188.2	34.0	595	14	CA825561 R61B04.TW
19	187.6	33.9	498	14	CF606656 GEMMA01.0
20	186.2	33.7	561	13	B0623572 USDA-FP-0
21	186.2	33.7	590	13	B0624319 USDA-FP-0
22	186.2	33.7	835	14	CB973441 CAB30003
23	186	33.6	729	14	CF518302 CAP0006.T
24	186	33.6	757	14	CB973510 CAP0006.T
25	186	33.6	769	14	CF519158 CAP0006.T
26	186	33.6	887	14	CF355583 V-B-129G1
27	185.2	33.5	765	14	CB347779 CAB2SG000
28	184.4	33.3	777	14	CF372668 GSECS079A
29	178.2	32.2	717	14	CF518533 CAP0007.T
30	175.8	31.8	747	14	CF231529 PAC0022B
31	173	31.3	666	13	BQ121429 EST607005
32	169.6	30.7	671	13	BQ414824 GA_Ed009
33	169.6	30.7	671	13	BQ406165 GA_Ed009
34	165	29.8	698	14	CF518449 CAP0007.T
35	158.2	28.6	571	14	CB347637 CAB2SG000
36	153.8	27.8	549	10	BE037501 MF02D05.M
37	153.8	27.8	549	12	BG524880 8-57.Stev
38	151.2	27.3	634	13	BQ991598 QGF23E22
39	146.6	26.5	749	14	CD846972 DHOA53ZA
40	145.8	26.4	391	9	AI443193 sa45f06.Y
41	144.6	26.1	552	14	CB347551 CAB2SG000
42	142.6	25.8	597	12	B1405821 146E08.MA
43	142.6	25.8	618	13	BQ121428 EST607004
44	140.8	25.5	144	14	CD413174 GM_CK4458
45	138.6	25.1	658	13	BQ987454 QGF12E07

## ALIGNMENTS

RESULT 1  
AM349814  
LOCUS  
DEFINITION  
AM349814  
ACCESSION  
AM349814.1 GI:6847453  
VERSION  
AM349814.1  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

1 (bases 1 to 645)  
Vodkin, L., Keim, P., Shoemaker, R., Retzel, E., Khanna, A., Corryell, V.,  
Expedding, J., Raph, C., Shoop, E., Pardinas, J., Liu, L., and Lewin, H.  
A Functional Genomics Program for Soybean (NSF 9872565)  
Unpublished (1999)  
Other ESTs: A144167  
Contact: Vodkin, L.O., PI, A Functional Genomics Program for  
Soybean (NSF 9872565)  
Lewin, H. A., Director, Keck Center for Comparative and Functional  
Genomics  
University of Illinois  
Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA  
Tel: (217) 244-6147

Fax: (217) 333-4582  
 Email: l-vodkin@uiuc.edu  
 This clone is available through: Genome Systems, Inc. 4633 World  
 Parkway Circle St. Louis, Missouri 63134. For further information  
 call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)  
 427-3324 or contact: clones@genomesystems.com or info@genome  
 systems.com web site: www.genomesystems.com  
 Seq primer: 5'-TTTATTTTATTTTATTTT(A/C/G)-3',  
 Location/Qualifiers

## FEATURES

source

1.645  
 /organism="Glycine max"  
 /mol\_type="mRNA"  
 /cultivar="Williams"  
 /db\_xref="taxon:3847"  
 /clone="Gm-r1021-1711"  
 /tissue\_type="root"  
 /lab\_host="X110-Gold"  
 /clone\_lib="Gm-r1021"  
 /note="Vector: Bluescript II XR, Site 1: EcoRI, Site 2:  
 XhoI; Library Gm-r1021 is a sequence-driven, resealed set  
 of the original library Gm-cl004 which was prepared from  
 root cDNA. The mRNA was isolated from entire roots of 8  
 day old 'Williams' seedlings which were propagated on  
 paper towels with distilled water. Stragene's cDNA  
 synthesis kit (catalog #200401) was used to synthesize  
 the cDNA. The Gm-cl004 library was constructed by Dr.  
 Paul Keim & Virginia H. Coryell, Department of Biology,  
 Box5640, Northern Arizona University, Flagstaff, AZ  
 86011, email: paul.keim@uau.edu, virginia.coryell@uau.edu.  
 The contig analysis to select unique genes was performed  
 by the laboratory of Ernest Retzel, Computational Biology  
 Centers, University of Minnesota,  
 http://www.cbc.umn.edu/ResearchProjects/Soybean/index.html  
 . Reracking was performed by Genome Systems, St. Louis,  
 http://www.genomesystems.com, and sequencing by the Keck  
 Center for Comparative and Functional Genomics,  
 University of Illinois,  
 http://www.life.uiuc.edu/biotech/keck.html."

## ORIGIN

Query Match 67.2%; Score 371.6; DB 10; Length 645;  
 Best Local Similarity 93.5%; Pred. No. 8.4e-90;  
 Matches 400; Conservative 0; Mismatches 24; Indels 4; Gaps 1;

QY 1 GTAGGTTATTTCCAGACATTAACACAACTTATTTGGGAGAGATTGTACTAGAAATAC 60  
 DB 88 GTAGTTTGGTCCAGACATTAACATTAATTATTTGGGAGAGATTGTACTAGAAATAC 147  
 QY 61 AAGTAGCGACTAAATTAACAAAACCTGAGGCTTTTGTACAGATAGCAATACCTCTC 120  
 DB 148 AAGTACCGACTCAAAATTAACAAAACCTGAGGCTTTTGTACAGAAATACAACTCTC 207  
 QY 121 TTTTATCATATCATATCTGAGTCTGCTGTTGGAATTTGACAGCACTTTTCCGAAACA 180  
 DB 208 TT-----TCTATCATATCATATCTGAGTCTGCTGTTGGAATTTGACAGCACTTTTCCGAAACA 263  
 QY 181 GGCTGTAATAAATAGCTTCAAGTCTGTCGCGCATCTGCTGCTTGTCTTGATGCAAGATATCC 240  
 DB 264 GGCTGTAATAAATAGCTTCAAGTCTGTCGCGCATCTGCTGCTTGTCTTGATGCAAGATATCC 323  
 QY 241 CTTCGTTCACTAGTCTCACTCTCCAGCTGCAAGTCACTAGTTCACCAATAGAACTCT 300  
 DB 324 CTTCGTTCACTAGTCTCACTCTCCAGCTGCAAGTCACTAGTTCACCAATAGAACTCT 383  
 QY 301 GGGAACATCATATGAGACCCCAATTAACCTCTCTCTCCCTGTATATGATCCATCCCTTTG 360  
 DB 384 GGGAACATCATATGAGACCCCAATTAACCTCTCTCTCTCCCTGTATATGATCCATCCCTTTG 443  
 QY 361 ACTTTTAAGGCTCACTCATATGATATGACCAACAGCTTTCTTTCGGAATATCAACG 420  
 DB 444 ACTTTTAAGGCTCACTCATATGATATGACCAACAGCTTTCTTTCGGAATATCAACG 503  
 QY 421 AATGCAAC 428

Db 504 AATGCAATC 511

RESULT 2  
 BM270700/c  
 LOCUS  
 DEFINITION  
 saki6e09.y1 Gm-cl075 Glycine max cDNA clone SOYBEAN CLONE ID:  
 Gm-cl075-3282 5', mRNA sequence.  
 ACCESSION  
 BM270700  
 VERSION  
 BM270700.1 GI:11763957  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Glycine max (soybean)  
 Glycine max  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.

REFERENCE  
 1 (bases 1 to 425)  
 Shoemaker R., Keim P., Vodkin L., Erpelting J., Coryell V.,  
 Khanna A., Bolla B., Marra M., Hillier L., Kucaba T., Martin J.,  
 Beck C., Wylie T., Underwood K., Steptoe M., Theising B., Allen M.,  
 Bowers J., Peterson B., Swaller T., Gibbons W., Pape D., Harvey N.,  
 Schurr R., Ritten B., Kohn S., Shin T., Jackson Y., Cardenas M.,  
 McCann R., Waterston R. and Wilson R.  
 Public Soybean EST Project  
 Unpublished (1999)  
 CONTACT: Shoemaker R./Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@wustl.edu

TITLE  
 JOURNAL  
 COMMENT  
 This clone is available through: Resgen, Invitrogen Corp. 2130  
 South Memorial Parkway Huntsville, AL 35801 For further information  
 call: (800)-533-4363 or contact: cduresgen.com web site:  
 www.resgen.com  
 Seq primer: -40RP from Gibco

High quality sequence scop: 413.

## FEATURES

source

1.425  
 Location/Qualifiers  
 /organism="Glycine max"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3847"  
 /clone="SOYBEAN CLONE ID: Gm-cl075-3282"  
 /tissue\_type="differentiating somatic embryos cultured on  
 MSM6AC"  
 /lab\_host="DH10B"  
 /clone\_lib="Gm-cl075"  
 /note="Vector: pBluescript II SK+; Site 1: EcoRI, Site 2:  
 XhoI. The cDNA library was constructed from mRNA isolated  
 from differentiating somatic embryos cultured on MSM6AC.  
 The library was prepared using the Stratagene pBluescript  
 II SK(+) library construction kit. Complementary DNA was  
 synthesized from mRNA using a primer consisting of a  
 poly(dT) sequence with an XhoI restriction site. EcoRI  
 adaptors were ligated to the blunt-ended cDNA fragments  
 followed by XhoI digestion. The cDNA fragments were  
 directionally cloned into the EcoRI-XhoI restriction site  
 of the pBluescript vector. The ligated cDNA fragments  
 were transformed into E. coli Electromax DH10B host cells.  
 Tissue culture and library construction were performed by  
 Francoise Thibaud-Nissen and Ann Khana (Lila Vodkin lab,  
 University of Illinois)."

## ORIGIN

Query Match 58.6%; Score 323.8; DB 12; Length 425;  
 Best Local Similarity 99.4%; Pred. No. 7.1e-77;  
 Matches 325; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTAGGTTATTTCCAGACATTAACAACTTATTTGGGAGAGATTGTACTAGAAATAC 60

Db	327	GTAGGTTAGATCCAGACATTAACAACAACCTTTATTTGGGAGAGATTGACTAGAAATAC	268
QY	61	AAGTGGCAGCTAAATTAATAAAAACGTCATGGCTTTTGTGCACAGATACGAAATACCTCCTC	120
Db	267	AAGTGGCAGCTAAATTAATAAAAACGTCATGGCTTTTGTGCACAGATACGAAATACCTCCTC	208
QY	121	TTTCATCTATCTATCTATCTCTGAGTTCCTGTTCGAATTGACGAACTTTTCCGACACA	180
Db	207	TTTCATCTATCTATCTATCTCTGAGTTCCTGTTCGAATTGACGAACTTTTCCGACACA	148
QY	181	GGCGTAAAAATAGCTTCAAGTCGTGGCCATCTGTGTCTTGTCTTGATGCAAAGATATCC	240
Db	147	GGCGTAAAAATAGCTTCAAGTCGTGGCCATCTGTGTCTTGTCTTGATGCAAAGATATCC	88
QY	241	CTTGCTTACTCAGTCTCACTTCCACCTGCAAGTCATCTAGTTCACCAATAGAACTCT	300
Db	87	CTTGCTTACTCAGTCTCACTTCCACCTGCAAGTCATCTAGTTCACCAATAGAACTCT	28
QY	301	GGGACATCTATATGACCCCCCAATTAAC	327
Db	27	GGGACATCTATATGACCCCCCAATTAAC 1	

RESULT 3				
BQ785857/c				
LOCUS	BQ785857	513 bp	mRNA	linear
DEFINITION	sa88602.y1 Gm-cl076 Glycine max cDNA clone SOYBEAN CLONE ID: Gm-cl076-6748 5', mRNA sequence.			
ACCESSION				
DESCRIPTION				

ACCESSION	BQ785857	GI:219943329
VERSION	BQ785857.1	
KEYWORDS	EST.	
SOURCE	Glycine max (soybean)	
ORGANISM	Glycine max	

SOURCE ORGANISM	REFERENCE AUTHORS
Glycine max (soybean)	
Glycine max	
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.	
1. (bases 1 to 513)	
	Shoenmaker, R., Keim, P., Vodkin, L., Eppelding, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, D., Kucaba, T., Mattlin, J., Beck, C., Wylie, T., Underwood, C., Steptoe, M., Thelsing, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.

**TITLE**  
 Public Soybean EST Project  
**JOURNAL**  
 Unpublished (1999)  
**COMMENT**  
 Contact: Shoemaker R/Public Soybean EST Project  
 Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available through: ResGen, Invitrogen Corp. 2130  
 South Memorial Parkway Huntsville, AL 35801 For further information  
 call: (800)-533-4363 or contact: [cu@resgen.com](mailto:cu@resgen.com) web site:  
[www.resgen.com](http://www.resgen.com)  
 Seq primer: -40RP from Gibco  
 High quality sequence, stop: 426.

```

FEATURES             Location/Qualifiers
     source            1..513
                        1. .513
                        /organism="Glycine max"
                        /mol_type="mRNA"
                        /db_xref="taxon:3847"
                        /clone="SOYBRAN CLONE ID: Gm-cl076-6748"
                        /tissue_type="wounded cotyledons"
                        /dev_stage="11 day old seedlings"
                        /lab_host="DH10B"
                        /clone_1fb="Gm-cl076"
                        /note="Vector: pluescript II SK+; Site_1: EcoRI; Site_2:
XhoI; The cDNA library was constructed from mRNA isolated
from 11 day old seedlings treated with that were treated
with 2 ugs/ml of a crude glucan elicitor preparation

```

isolated from the mycelial walls of *Streptophthora soiae*. The library was prepared using the Stratagene Bluescript II SK(+) library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with an XhoI restriction site. EcoRI adaptors were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the Bluescript vector. The ligated cDNA fragments were transformed into *E. coli* Electromax DH10B host cells. Plant material was provided by Michael G. Hahn (Complex Carbohydrate Research Center, University of Georgia) and the library was constructed by Anu Khanna (Ulla Vedkin lab, University of Illinois)."

Query Match	55.2%;	Score 305.4;	DB 13;	length 513;
Best Local Similarity	93.0%;			
Matches 332;	Conservative	0;	Mismatches 21;	Indels 4;
				Gaps 1;

QY	1	GTAGTATTATTCAGACATTACACACACTTATTTCGGAGAGAGATTGTACTGAATAAC	60
Db	409	GTAGTTTGGTCAGACATTAACTAACTTAACTTTTCGGAGAGAGATTGTACTGAATAAC	350
QY	61	AAGAAGGACTAAATAATACAAAACTGCAGAGCTTTTGTACAGATACGAATCTTCTC	120
Db	349	AAGTACCGACTCAATATACAAAACTGCAGAGTTTTTGTACAAATACGAAAACTACTTC	290
QY	121	TTTCTATCTATCTATCTATCTCTGAGTCTCTGTTCGAATTCGACAACTTTCCCGACA	180
Db	289	TT-----TCATCTATCTATCTTTGAGTTCCTGTTCAAATTGAAGCAACTTTCCCGACA	234
QY	181	GGCTGTAAAAATAGCTTCAAGTGGTTGGCATCTGTGTCTTGTCTTGATGCAAGATATCC	240
Db	233	GGCTGTAAAAATAGCTTCAAGTGGTGAATGTGTCTTGTCTTGATGCAAGATATCC	174
QY	241	CTTGTCTTCACTCACTCTCACTTCCACCTCGCAATCTATAGTTCACCAATGGAATCTC	300
Db	173	CTTGTCTTCACTCACTCTCACTTCAACTGGAATTCACTTCACTTCAACCAATGGAATCTC	114
QY	301	GGGACATCTATATAGACCCCAATAAACTTCTCTCTCTTGATGATCATTTCCCTC	357
Db	113	GGGACATCTATATAGACCCCAAGAACTTCTCTCTCTTGATGATCATTTCCCTC	57

RESULT 4	383 bp	mRNA	linear	EST 28-NOV-2001
LOCUS	BG238159/c			
DEFINITION	BG238159			
DESCRIPTION	sabb6h12.y1-Gm-cl043 Glycine max cDNA clone GENOME SYSTEMS CLONE			
ACCESSION	ID: Gm-11043-4199 5', mRNA sequence.			
VERSION	BG238159			
KEYWORDS	BG238159.1	GI:12773232		
SOURCE	EST			
ORGANISM	Glycine max (soybean)			

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 383)	Shoemaker, R., Keim, P., Vodka, L., Erpelidg, J., Coyvell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Gleason, M., Theising, B., Allen, M., Bowers, Y., Peterson, B., Swaller, T., Stibbons, M., Pape, D., Harvey, N., Schmitt, R., Ritter, E., Kohn, S., Shih, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.	Public Soybean EST Project	Unpublished (1999)	Contact: Shoemaker R./Public Soybean EST Project Public Soybean EST Project Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800

Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available through: Resgen, Invitrogen Corp. 2130  
South Memorial Parkway Huntville, AL 35801 For further information  
call: (800)-533-4363 or contact via email: cw@resgen.com  
High quality sequence stop: 366.  
Location/Qualifiers

## FEATURES

source

1..383  
/organism="Glycine max"  
/mol\_type="mRNA"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-c1043-4199"  
/issue\_type="Hypocotyl and Plumule, germinating seeds"  
/lab\_host="DH10B"  
/clone\_lib="Gm-c1043"  
/note="Vector: pT73Pac (Pharmacia); Site\_1: EcoRI;  
Site\_2: NotI; This cDNA library was constructed from mRNA  
isolated from hypocotyl and plumule tissues of seeds  
germinated for three days of the cultivar Williams.  
Complementary DNA was synthesized from mRNA using a primer  
consisting of a poly(dT) sequence with a NotI restriction  
site. EcoRI adapters were ligated to the blunt-ended cDNA  
fragments followed by digestion with EcoRI and NotI. The  
cDNA fragments were directionally cloned into the  
EcoRI-NotI restriction site of the pT73-Pac vector. The  
ligated cDNA fragments were transformed into DH10B host  
cells (Gibco BRL). This library was constructed by Dr.  
Randy Shoemaker."

## ORIGIN

Query Match 54.4%; Score 300.8; DB 12; Length 383;  
Best Local Similarity 99.3%; Pred. No. 1.2e-70;  
Matches 302; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTAGCTATTTCGCGACATTACCACTTATTTCGGGAAGATTGTACTAGAAATAC 60  
Db 304 GTAGCTATTTCGCGACATTACCACTTATTTCGGGAAGATTGTACTAGAAATAC 245  
QY 61 AAGTAGGAGCTAAATATACAAAACATGCGCTTTTGTGCAGATGAGATATCTTCCTC 120  
Db 244 AAGTAGGAGCTAAATATACAAAACATGCGCTTTTGTGCAGATGAGATATCTTCCTC 185  
QY 121 TTTCATCTATCTATCTATCTCTGAGTTCCTGTTGCAATGCGCAACTTTCCGGAACA 180  
Db 184 TTTCATCTATCTATCTATCTCTGAGTTCCTGTTGCAATGCGCAACTTTCCGGAACA 125  
QY 181 GCGCTGAATAATACCTCAAGCTGTTGGCACTGTGCTGTTGATGCAAGATATCC 240  
Db 124 GCGCTGAATAATACCTCAAGCTGTTGGCACTGTGCTGTTGATGCAAGATATCC 65  
QY 241 CTGCTCTCACTCACTCTCACTCCAGCTGACATCTAGTTCACCAATAGAGACTCT 300  
Db 64 CTGCTCTCACTCACTCTCACTCCAGCTGACATCTAGTTCACCAATAGAGACTCT 5  
QY 301 GGGA 304  
Db 4 GGGA 1

RESULT 5  
AM620829/c 432 bp mRNA linear EST 03-DEC-2001  
LOCUS s147a12.y1 Gm-c1033 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:  
DEFINITION Gm-c1033-407 5', mRNA sequence.  
ACCESSION AM620829  
VERSION AM620829.1 GI:7327031  
KEYWORDS EST.  
SOURCE Glycine max (soybean)  
ORGANISM Glycine max  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.

## REFERENCE

AUTHORS

1 (bases 1 to 432)  
Shoemaker, R., Keim, P., Vodkin, L., Erpelidng, J., Coryell, V.,  
Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,  
Beck, C., Wylie, T., Underwood, K., Seepoe, M., Theising, B., Allen, M.,  
Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,  
Schurk, R., Ritter, E., Kohn, S., Shit, T., Jackson, Y., Cardenas, M.,  
McCam, R., Waterston, R. and Wilson, R.  
Public Soybean EST Project  
Unpublished (1999)  
Contact: Shoemaker R/Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@waton.wustl.edu

COMMENT

This clone is available through: Resgen, Invitrogen Corp. 2130  
South Memorial Parkway Huntville, AL 35801 For further information  
call: (800)-533-4363 or contact via email: cw@resgen.com  
Insert Length: 649 Std Error: 0.00  
High quality sequence stop: 414.  
Location/Qualifiers

## FEATURES

source

1..432  
/organism="Glycine max"  
/mol\_type="mRNA"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-c1033-407"  
/issue\_type="Desloy 5710' seedling roots"  
/lab\_host="DH10B"  
/clone\_lib="Gm-c1033"  
/note="Vector: pBluescript II XR; Site\_1: EcoRI; Site\_2:  
XhoI; This cDNA library was constructed from mRNA isolated  
from 'Desloy 5710' seedling roots. Tissue was taken from  
7-day-old seedlings that had been propagated on paper  
towels with distilled water. Tissue was taken from the tip  
to the first lateral root, usually about 3cm from the tip,  
and flash-frozen in liquid nitrogen. StrataGene's cDNA  
Synthesis Kit (catalog number 200401) was used to  
synthesize the cDNA. First-strand synthesis was performed  
with 5-methyl dCTP, hence the ligated cDNA was  
hemimethylated. A modification of StrataGene's  
first-strand synthesis primer was used. An 'anchor'  
nucleotide (V=A, C, or G) was added to the 3' end of the  
primer [GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG(T)18V] to anchor  
the primer at the 5' end of the poly(A) tract. After  
second strand synthesis, the cDNA ends were filled in with  
cloned Pfu DNA, ligated to EcoRI adapters and  
subsequently phosphorylated. The cDNA was then  
precipitated and redissolved in sterile, RNase-  
DNase-free water. The XhoI site within the first-strand  
synthesis primer was then restricted by digestion with  
XhoI from Promega (400U/ul); all XhoI sites in the cDNA  
would be protected by their hemimethylated status. The  
cDNA constructs were size-fractionated with a 500 bp  
cutoff, using GabcORL Life Technologies' cDNA size  
fractionation column. The column eluent was then  
precipitated, redissolved, and ligated into StrataGene's  
pBluescript II XR predigested vector (pBluescript II SK(+)  
vector that has been digested with EcoRI and XhoI, and  
phosphorylated by StrataGene). 100% of the white and blue  
colonies appear to contain recombinant plasmids with cDNA  
inserts, based on size (n=24 and 6, respectively). This  
library was constructed by Dr. Paul Keim and Dr. Virginia  
Coryell."

## ORIGIN

Query Match 50.5%; Score 279.2; DB 10; Length 432;  
Best Local Similarity 96.5%; Pred. No. 8.5e-65;  
Matches 299; Conservative 0; Mismatches 3; Indels 8; Gaps 1;

QY 127 TCATCTATCTATCTCTGAGTTCCTGTTGCAATGAGCAACTTTCCGGAAGAGCTGT 186  
Db 432 TCATCTATCTATCTCTGAGTTCCTGTTGCAATGAGCAACTTTCCGGAAGAGCTGT 373

QY 187 AAAATGCTTCAAGTCGTTCCGATCTGTCTTGTCTTGATGCAAGATATCCCTTGT 246  
 Db 372 AAACATAGCTTCAAGTCGTTCCGATCTGTCTTGTCTTGATGCAAGATATCCCTTGT 313  
 QY 247 TCAGTCAAGTCGTTCCGATCTGTCTTGTCTTGATGCAAGATATCCCTTGT 306  
 Db 312 TCAGTCAAGTCGTTCCGATCTGTCTTGTCTTGATGCAAGATATCCCTTGT 253  
 QY 307 TCATATGACCCCAATTAACCTTCTCTCTCTTGATGATTCATTC-----CCCTT 358  
 Db 252 TCATATGACCCCAATTAACCTTCTCTCTCTTGATGATTCATTC-----CCCTT 193  
 QY 359 TCAGTCAAGTCGTTCCGATCTGTCTTGTCTTGATGCAAGATATCCCTTGT 418  
 Db 192 TCAGTCAAGTCGTTCCGATCTGTCTTGTCTTGATGCAAGATATCCCTTGT 133  
 QY 419 TGAATGCAAC 428  
 Db 132 TGAATGCAAC 123

RESULT 6 BU964888 375 bp mRNA linear EST 21-OCT-2002  
 LOCUS sat03h07.y1 Gm-c1036 Glycine max cDNA clone SOYBEAN CLONE ID:  
 DEFINITION BU964888

ACCESSION BU964888.1 GI:24205635  
 VERSION EST.  
 KEYWORDS Glycine max (soybean)  
 SOURCE Glycine max  
 ORGANISM Glycine max (soybean)

REFERENCE 1 (bases 1 to 375)  
 AUTHORS Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Corryell, V.,  
 Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,  
 Beck, C., Wylie, T., Underwood, K., Stepien, M., Theising, B., Allen, M.,  
 Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,  
 Schurr, R., Ritzer, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,  
 McGinn, R., Waterston, R. and Wilson, R.  
 TITLE Public Soybean EST Project  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Shoemaker R/Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available through: Reggen, Invitrogen Corp. 2130  
 South Memorial Parkway Huntville, AL 35801 For further information  
 call: (800)-533-4363 or contact: cdueregen.com web site:  
 www.reggen.com  
 Putative full length read  
 Vector to vector length is 376  
 Seq primer: -40RP from Gibco.

## FEATURES

source  
 1..375  
 /organism="Glycine max"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:3847"  
 /clone="SOYBEAN CLONE ID: Gm-c1036-12662"  
 /tissue\_type="somatic embryos cultured on MSD 20"  
 /lab\_host="DH10B"  
 /clone\_id="Gm-c1036"  
 /note="Vector: pSPORT1, Site 1: NotI, Site 2: SalI. This  
 cDNA library was constructed from mRNA isolated from  
 somatic embryos (age ranging from 2 months to 9 months)  
 cultured on MSD 20. The library was prepared using the  
 Life Technologies superscript cDNA library construction

## ORIGIN

Query Match 49.3%; Score 272.8; DB 13; Length 375;  
 Best Local Similarity 99.3%; Pred. No. 4.5e-63;  
 Matches 274; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

kit. Complementary DNA was synthesized from mRNA using a  
 poly (dt) sequence with a NotI restriction site. SalI  
 linkers adapters were ligated to the blunt-ended cDNA  
 fragments followed by NotI digestion. The cDNA fragments  
 were directionally cloned into the NotI-SalI restriction  
 site of the pSPORT1 vector. The ligated cDNA fragments  
 were transformed into E. coli Electromax DH10B host cells.  
 This library was constructed in the laboratory of Dr. Lila  
 Vodkin by Anu Khanna at the University of Illinois at  
 Urbana-Champaign. e-mail: l.vodkin@uiuc.edu"

QY 1 GTAGTATATTTCCAGCATTAACACAACTTTATTTCCGGAAGATTTGACTGAAATAC 60  
 Db 276 GTAGTATATTTCCAGCATTAACACAACTTTATTTCCGGAAGATTTGACTGAAATAC 217  
 QY 61 AAGTAGCCACTAAATTAACAAAGCTGATGCTTTGTGACAGATAGCATTAATCTCC 120  
 Db 216 AAGTAGCCACTAAATTAACAAAGCTGATGCTTTGTGACAGATAGCATTAATCTCC 157  
 QY 121 TTTTATCTATCTATCTATCTGATGCTTCTGTTGATGCAAGATTTCCGGAACA 180  
 Db 156 TTTTATCTATCTATCTATCTGATGCTTCTGTTGATGCAAGATTTCCGGAACA 97  
 QY 181 GCGCTGTAATAAATGCTTCAAGTCGTTGGGCACTGTCTTGTCTTGATGCAAGATATCC 240  
 Db 96 GCGCTGTAATAAATGCTTCAAGTCGTTGGGCACTGTCTTGTCTTGATGCAAGATATCC 37  
 QY 241 CTGGCTCACTCAAGTCCTCACTCCAGCTGCAAGTCA 276  
 Db 36 CTGGCTCACTCAAGTCCTCACTCCAGCTGCAAGTCA 1

RESULT 7 BH405680 576 bp DNA linear GSS 04-DEC-2001  
 LOCUS Gm-UMB001\_109\_J13 R UMN Soybean BAC Library (pCISBAC ECKR1)  
 DEFINITION Glycine max genomic clone Glycine max genomic clone  
 Gm-UMB001\_109\_J13, genomic survey sequence.  
 ACCESSION BH405680.1 GI:17351835  
 VERSION GSS.  
 KEYWORDS Glycine max (soybean)  
 SOURCE Glycine max  
 ORGANISM Glycine max (soybean)  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 rosids; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.

REFERENCE 1 (bases 1 to 576)  
 AUTHORS Marek, J.F., Paz, M., Darnielle, L., Hanson, N. and Shoemaker, R.C.  
 TITLE BAC End sequences from a soybean genomic library (ISU)  
 JOURNAL Unpublished (2000)  
 COMMENT Contact: Shoemaker Randy C  
 Agronomy Department  
 Iowa State University  
 Ames, IA 50011-1010, USA  
 Tel: 515 294 1205  
 Fax: 515 294 2299  
 Email: rcshoe@iastate.edu  
 This BAC identified by SSR Sat2429. For more information, see  
 SoyBase at:  
 http://genome.cornell.edu/cgi-bin/webace/webace?db=soybase. Please  
 see as an authority for the mapping/finishing: Cregan P.B., T. Jarvik,  
 A.L. Bush, R.C. Shoemaker, K.G. Laik, A.L. Kahler, N. Kaya, T.T.  
 Vantrolat, D.G. Lohnes, J. Chung, and J.E. Specht. 1999a. An  
 integrated genetic linkage map of the soybean genome. Crop Sci.  
 39:1464-1490.

Insert Length: 125 Std Error: 0.00  
 Seq primer: M13R

Class: BAC ends.  
Location/Qualifiers  
source  
1.576

/organism="Glycine max"  
/mol\_type="genomic DNA"  
/cultivar="Fairdault"  
/db\_xref="taxon:3847"  
/clone="Gm UMD001.109 J13"  
/issue\_type="cotyledon leaves"  
/dev\_stage="cotyledon"  
/clone\_1lb="UMD Soybean BAC library (pECsBAC4 EcORI)  
Glycine max genomic clone"  
/note="Vector: pECsBAC4; The UMD BAC library (Danesh et al, Theor. Appl. Genet. 96:196, 1998) was constructed using the Eco RI site of pECsBAC4. The library consists of 72,960 clones with an average insert size of 120 kb, equal to 7 haploid genome equivalents. Screening of the library is done by hybridization of high-density colony filters and/or PCR amplification of DNA pools. Four high density filters, each containing 18,432 clones (doubly spotted), represent the whole library for colony screening"

## ORIGIN

Query Match 46.3%; Score 255.8; DB 28; Length 576;  
Best Local Similarity 97.3%; Pred. No. 1.9e-58;  
Matches 257; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 GTAGGTTATTTCCAGCATTAACAACAATTATTTGGGAGAGATTGTACTGAAATAC 60  
DB 285 GTAGGTTAATCCAGCATTAACAACAATTATTTGGGAGAGATTGTACTGAAATAC 344  
QY 61 AAGTACGCACTAAATTAACAACAAGTGGCTTTGTGTCAGATGCAATCTCTCC 120  
DB 345 AAGTACGCACTAAATTAACAACAAGTGGCTTTGTGTCAGATGCAATCTCTCC 404  
QY 121 TTTCTATCTATCTATCTATCTGAGTCTGTTGCAATGACGAACCTTTCCGAGA 180  
DB 405 TTTCTATCTATCTATCTATCTGAGTCTGTTGCAATGACGAACCTTTCCGAGA 464  
QY 181 GGCTGTAATAATGCTTCAAGTCGTGGCATCTGTCTGTCTTGATGCAAGATATCC 240  
DB 465 GGCTGTAATAATGCTTCAAGTCGTGGCATCTGTCTGTCTTGATGCAAGATATCC 524  
QY 241 CTGGCTTCACTAGTCTCACTTCC 264  
DB 525 CTGGCTTCACTAGTCTCACTTCC 548

RESULT 8  
BE943027/c 619 bp mRNA linear EST 03-OCT-2000  
LOCUS EST422606 MGHG Medicago truncatula cDNA clone pmGHG-14H1, mRNA  
DEFINITION  
sequence.  
ACCESSION BE943027  
VERSION BE943027.1 GI:10520786  
KEYWORDS EST.  
SOURCE Medicago truncatula (barrel medic)  
ORGANISM Medicago truncatula  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eustosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.

REFERENCE 1 (bases 1 to 619)  
AUTHORS Cote, F., Ojane-Reuthe, T., Hahn, M. G., Vandenberg, K., Hur, J., Beremand, P., Andre, G., Town, C. D., Bowman, C. L., Crahen, M. B. and Cho, J., Fraser, C. M.  
TITLE ESTs from seedling roots of Medicago truncatula after treatment with beta glucan elicitor preparation from Phytophthora sojae  
JOURNAL Unpublished (2000)  
COMMENT Contact: Michael G. Hahn  
Complex Carbohydrate Research Center  
University of Georgia

220 Riverbend Road, Athens, GA 30602-4712, USA  
Tel: 706-542-4457  
Fax: 706-542-4412  
Email: hahn@ccr.cuga.edu  
University of Georgia name: G270879e TIGR sequence name: MTUBE37TK  
More information is available at: <http://chryslie.tamu.edu/medicago>  
Seq primer: Skmod (CTA GAA CTA gtc gat cc).  
Location/Qualifiers  
1.619  
/organism="Medicago truncatula"  
/mol\_type="mRNA"  
/cultivar="A17"  
/db\_xref="taxon:3880"  
/clone="pmGHG-14H1"  
/issue\_type="Roots from four day old seedlings"  
/dev\_stage="2 days after treatment with beta glucan elicitor preparation from Phytophthora sojae"  
/lab\_host="E. coli strain XL0LR"  
/clone\_1lb="MGHG"  
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA was directionally ligated into the Unizap XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-Zap phage using Ex-assist helper phage and propagated in XL0LR cells."

## FEATURES

source

## ORIGIN

Query Match 41.3%; Score 228.2; DB 10; Length 619;  
Best Local Similarity 85.5%; Pred. No. 5.6e-51;  
Matches 254; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 132 CATATCATCTGAGTCTGCTGCTGCAATGACGAACCTTTCCGAGAGCTGTAAAA 191  
DB 389 CATATCATCTGAGTCTGCTGCTGCAATGACGAACCTTTCCGAGAGCTGTAAAA 330  
QY 192 TAGCTTCAAGTGGTGGCATCTGTCTGTTGTGATGCAAGATATCCCTGCTTCACT 251  
DB 329 TAGCTTCAAGTGGTGGCATCTGTCTGTTGTGATGCAAGATATCCCTGCTTCACT 270  
QY 252 CAGTCTCACTTCCAGCTGCACTGATCTAGTTCACCAATGAGAACTGTGGACATCTAT 311  
DB 269 AAGTCTCACTTCCAGCTGCACTGATCTAGTTCACCAATGAGAACTGTGGACATCTAT 210  
QY 312 AAGACCCCCCAATAAATCTTCTCTGCTGTAAGATATCCCTTGAAGCTTAAAGCT 371  
DB 209 AAGACCCCCCAATAAATCTTCTCTGCTGTAAGATATCCCTTGAAGCTTAAAGCT 150  
QY 372 CAATCATATGTATATACCAACAGTCTTCTGTTCCGAATATCACATGATGCAAC 428  
DB 149 CAATCATATGTATATACCAACAGTCTTCTGTTCCGAATATCACATGATGCAAC 93

RESULT 9  
CB891233/c 637 bp mRNA linear EST 24-APR-2003  
LOCUS EST448202 KV3 Medicago truncatula cDNA clone KV3-49H9, mRNA  
DEFINITION  
sequence.  
ACCESSION CB891233  
VERSION CB891233.1 GI:30099400  
KEYWORDS EST.  
SOURCE Medicago truncatula (barrel medic)  
ORGANISM Medicago truncatula  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eustosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.

REFERENCE 1 (bases 1 to 637)  
AUTHORS Vandenberg, K., Andre, G., Hur, J., Beremand, P., Town, C. D., Van Aken, S., Ullrich, T., Cho, J. and Fraser, C. M.  
TITLE ESTs from roots of Medicago truncatula 72 h after Rhizobium inoculation, 2001  
JOURNAL Unpublished (2001)

COMMENT Contact: Vandenberg K  
Department of Plant Biology  
University of Minnesota  
220 Biosci Center, 1445 Gortner Ave, St. Paul, MN 55108, USA  
Tel: 612 624 2755  
Fax: 612 625 1738  
Email: kvandenb@cs.umn.edu  
TIGR sequence name: MTECO41TK  
More information is available at: www.medicago.org  
Seq primer: SKmod (CTA GAA CTA gtc gat cc).  
Location/Qualifiers

FEATURES  
source  
1..637  
/organism="Medicago truncatula"  
/mol\_type="mRNA"  
/cultivar="genotype A17"  
/db\_xref="taxon:3880"  
/clone="KV3-49H9"  
/tissue\_type="Seedling roots"  
/dev\_stage="3 days post-inoculation with Sinorhizobium meliloti"  
/lab\_host="E. coli strain X10LR"  
/clone\_lib="KV3"  
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA was directionally ligated into the Unizap XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-Zap phage using Ex-assist helper phage and propagated in X10LR cells."

ORIGIN  
Query Match 41.3%; Score 228.2; DB 14; Length 637;  
Best Local Similarity 85.5%; Pred. No. 5.6e-51;  
Matches 254; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 132 CTATCTATCTCTGAGTCTCTGTCGAAATTCGAGCACTTTCCGAAACAGGCTGTAAAAA 191  
DB 632 CTATCTATCTCTGAGTCTCTGTCGAAATTCGAGCACTTTCCGAAACAGGCTGTAAAAA 573  
QY 192 TAGCTTCAAGTCGTGGGCACTGTGTCTGTCGTAATGAGATATCCGATGCTTCACT 251  
DB 572 TAGCTTCAAGTCGTGGGCACTGTGTCTGTCGTAATGAGATATCCGATGCTTCACT 513  
QY 252 CAGTCTCACTTCCAGCTGCAAGTCATGATGATGCAATGAGATCTGGGACATCTAT 311  
DB 512 AAGCTCACTTCCAGCTGCAAGTCATGATGATGCAATGAGATCTGGGACATCTAT 453  
QY 312 ATGACCCCAATAAATCTTCTCTCTGTCGTAATGATGATGATGATGATGATGAT 371  
DB 452 ATGACCCCAATAAATCTTCTCTCTGTCGTAATGATGATGATGATGATGATGAT 393  
QY 372 CAATCTATATGTAATACCAACAGCTTTCTGTCGAACTATACCACTGAATGCAAC 428  
DB 392 CAATCTATATGTAATACCAACAGCTTTCTGTCGAACTATACCACTGAATGCAAC 336

RESULT 10  
BM812845/c 680 bp mRNA linear EST 05-MAR-2002  
EST530938 BNIR Medicago truncatula/Meloidogyne incognita mixed EST  
library cDNA clone pBNIR-4E8, mRNA sequence.

ACCESSION BM812845  
VERSION BM812845.1 GI:19148859  
KEYWORDS EST.  
SOURCE Medicago truncatula/Meloidogyne incognita mixed EST library  
ORGANISM Medicago truncatula/Meloidogyne incognita mixed EST library  
Eukaryota; mixed EST libraries.  
REFERENCE 1 (bases 1 to 680)  
AUTHORS Bird,D., Kolitz,H., Samac,D., Town,C.D., Van Aken,S., Uteback,T.,  
Cheung,F., Tsai,J. and Fraser,C.M.  
TITLE ESTs from roots of Medicago truncatula after infection with the  
JOURNAL nematode Meloidogyne incognita  
Unpublished (2002)

COMMENT Contact: Bird, DM  
Plant Nematode Genetics Group  
North Carolina State University  
Box 7616, Raleigh, NC 27695, USA  
Tel: 919 515 6813  
Fax: 919 515 9500  
Email: david\_bird@ncsu.edu  
TIGR sequence name: MTQAB28TK More information is available at:  
www.medicago.org  
Seq primer: SKmod (CTA GAA CTA gtc gat cc).  
Location/Qualifiers

FEATURES  
source  
1..680  
/organism="Medicago truncatula/Meloidogyne incognita mixed  
EST library"  
/mol\_type="mRNA"  
/cultivar="A17"  
/db\_xref="taxon:188702"  
/clone="pBNIR-4E8"  
/tissue\_type="3 week old roots"  
/dev\_stage="3 days after infection with Meloidogyne incognita second stage larvae"  
/lab\_host="X10LR"  
/clone\_lib="BNIR"  
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA was directionally ligated into the Unizap XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-Zap phage using Ex-assist helper phage and propagated in X10LR cells."

ORIGIN  
Query Match 41.3%; Score 228.2; DB 12; Length 680;  
Best Local Similarity 85.5%; Pred. No. 5.6e-51;  
Matches 254; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 132 CTATCTATCTCTGAGTCTCTGTCGAAATTCGAGCACTTTCCGAAACAGGCTGTAAAAA 191  
DB 656 CTATCTATCTCTGAGTCTCTGTCGAAATTCGAGCACTTTCCGAAACAGGCTGTAAAAA 597  
QY 192 TAGCTTCAAGTCGTGGGCACTGTGTCTGTCGTAATGAGATATCCGATGCTTCACT 251  
DB 596 TAGCTTCAAGTCGTGGGCACTGTGTCTGTCGTAATGAGATATCCGATGCTTCACT 537  
QY 252 CAGTCTCACTTCCAGCTGCAAGTCATGATGATGCAATGAGATCTGGGACATCTAT 311  
DB 536 AAGCTCACTTCCAGCTGCAAGTCATGATGATGCAATGAGATCTGGGACATCTAT 477  
QY 312 ATGACCCCAATAAATCTTCTCTCTGTCGTAATGATGATGATGATGATGATGAT 371  
DB 476 ATGACCCCAATAAATCTTCTCTCTGTCGTAATGATGATGATGATGATGATGAT 417  
QY 372 CAATCTATATGTAATACCAACAGCTTTCTGTCGAACTATACCACTGAATGCAAC 428  
DB 416 CAATCTATATGTAATACCAACAGCTTTCTGTCGAACTATACCACTGAATGCAAC 360

RESULT 11  
BM814531/c 730 bp mRNA linear EST 05-MAR-2002  
EST52625 BNIR Medicago truncatula/Meloidogyne incognita mixed EST  
library cDNA clone pBNIR-22623, mRNA sequence.

ACCESSION BM814531  
VERSION BM814531.1 GI:19150545  
KEYWORDS EST.  
SOURCE Medicago truncatula/Meloidogyne incognita mixed EST library  
ORGANISM Medicago truncatula/Meloidogyne incognita mixed EST library  
Eukaryota; mixed EST libraries.  
REFERENCE 1 (bases 1 to 730)  
AUTHORS Bird,D., Kolitz,H., Samac,D., Town,C.D., Van Aken,S., Uteback,T.,  
Cheung,F., Tsai,J. and Fraser,C.M.  
TITLE ESTs from roots of Medicago truncatula after infection with the  
JOURNAL nematode Meloidogyne incognita



## JOURNAL

Unpublished (2002)

Other ESTs: ES1592624

Contact: Bird, DM

Plant Nematode Genetics Group

North Carolina State University

Box 7616, Raleigh, NC 27695, USA

Tel: 919 515 6813

Fax: 919 515 9500

Email: david.bird@ncsu.edu

TIGR sequence name: MT04U481KB More information is available at:

www.medicago.org

Seq primer: Skmod (CTA gga cta gtc gat cc).

## FEATURES

Source

Location/Qualifiers

1..730

/organism="Medicago truncatula/Meloidogyne incognita mixed

EST library"

/mol\_type="mRNA"

/cultivar="A17"

/db\_xref="taxon:186702"

/clone="pENIR-22G23"

/issue\_type="3 week old roots"

/dev\_stage="3 days after infection with Meloidogyne

incognita second stage larvae"

/lab\_host="XIOIR"

/clone\_lib="ENIR"

/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:

XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA

was directionally ligated into the UniZap XR vector from

Stratagene and packaged using Gigapack III Gold packaging

extracts. Plasmids containing cDNA inserts were excised

from the recombinant lambda-Zap phage using Ex-assist

helper phage and propagated in XIOIR cells."

## ORIGIN

Query Match

Best Local Similarity 41.3%; Score 228.2; DB 12; Length 730;

Matches 254; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 132 CTATCATCTCTGAGTCTCTGTTCCGATTCGACCACTTTCCGAAACAGGCTGTAAAAA 191

DB 697 CTATCATCTCTGAGTCTCTGTTCCGATTCGACCACTTTCCGAAACAGGCTGTAAAAA 638

QY 192 TAGCTTCAAGTGGTGGCGATCTGTCTGTTGATGCAAGATATCCCTTCTTCACT 251

DB 637 TAGCTTCAAGTGGTGGCGATCTGTCTGTTGATGCAAGATATCCCTTCTTCACT 578

QY 252 CAGTTCACCTTCACCTGCAAGTCATCTAGTTCACCAATGAGAACTCTGGACATCTAT 311

DB 577 AAGTTCACCTTCACCTGCAAGTCATCTAGTTCACCAATGAGAACTCTGGACATCTAT 518

QY 312 ATGACCCCCCAATAACTTTCTCTCTCTGTATGATTCATTCCTCTTGAATTTAAGGT 371

DB 517 ATGACCCCCCAATAACTTTCTCTCTCTGTATGATTCATTCCTCTTGAATTTAAGGT 458

QY 372 CAAGTCATATGTATATACCAACAGTCTCTGTTCCGAACTATACCATGATGCAAC 428

DB 457 CAAGTCATATGTATATACCAACAGTCTCTGTTCCGAACTATACCATGATGCAAC 401

RESULT 12

LOCUS CA922589 758 bp mRNA linear EST 09-MAY-2003

DEFINITION EST640307 MTUS Medicago truncatula cDNA clone MTUS-55G1, mRNA

sequence.

ACCESSION CA922589

VERSION CA922589.1 GI:27409519

KEYWORDS EST.

SOURCE Medicago truncatula (barrel medic)

ORGANISM Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;

Medicago.

## REFERENCE

1 (bases 1 to 758)

Vandenbosch, K., Endre, G., Silverstein, K., Town, C.D., Van Aken, S.,

Utterback, T., Cheung, F., and Fraser, C.M.

The Medicago truncatula 6X unigene set: cDNA clones selected and

re-arranged from various libraries

Unpublished (2002)

Contact: Vandenbosch K

Department of Plant Biology

University of Minnesota

220 BiocSci Center, 1445 Gortner Ave, St. Paul, MN 55108, USA

Tel: 612 624 2755

Fax: 612 625 1738

Email: kvandenb@cbbs.umn.edu

Alias Clone pMRAM-47E20

TIGR sequence name: MTUCC731V

More information is available at: www.medicago.org

Seq primer: (gca ata cga ctc act ata ggg c).

## FEATURES

Source

Location/Qualifiers

1..758

/organism="Medicago truncatula"

/mol\_type="mRNA"

/cultivar="A17"

/db\_xref="taxon:3880"

/clone="MTUS-55G1"

/issue\_type="mixed tissues"

/dev\_stage="various stages"

/lab\_host="XIOIR"

/clone\_lib="MTUS"

/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:

XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA

was directionally ligated into the UniZap XR vector from

Stratagene and packaged using Gigapack III Gold packaging

extracts. Plasmids containing cDNA inserts were excised

from the recombinant lambda-Zap phage using Ex-assist

helper phage and propagated in XIOIR cells."

## ORIGIN

Query Match

Best Local Similarity 41.3%; Score 228.2; DB 14; Length 758;

Matches 254; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 132 CTATCATCTCTGAGTCTCTGTTCCGATTCGACCACTTTCCGAAACAGGCTGTAAAAA 191

DB 448 CTATCATCTCTGAGTCTCTGTTCCGATTCGACCACTTTCCGAAACAGGCTGTAAAAA 507

QY 192 TAGCTTCAAGTGGTGGCGATCTGTCTGTTGATGCAAGATATCCCTTCTTCACT 251

DB 508 TAGCTTCAAGTGGTGGCGATCTGTCTGTTGATGCAAGATATCCCTTCTTCACT 567

QY 252 CAGTTCACCTTCACCTGCAAGTCATCTAGTTCACCAATGAGAACTCTGGACATCTAT 311

DB 568 AAGTTCACCTTCACCTGCAAGTCATCTAGTTCACCAATGAGAACTCTGGACATCTAT 627

QY 312 ATGACCCCCCAATAACTTTCTCTCTCTGTATGATTCATTCCTCTTGAATTTAAGGT 371

DB 628 ATGACCCCCCAATAACTTTCTCTCTCTGTATGATTCATTCCTCTTGAATTTAAGGT 687

QY 372 CAAGTCATATGTATATACCAACAGTCTCTGTTCCGAACTATACCATGATGCAAC 428

DB 688 CAAGTCATATGTATATACCAACAGTCTCTGTTCCGAACTATACCATGATGCAAC 744

RESULT 13

LOCUS AQ989424 532 bp DNA linear GSS 06-MAR-2000

DEFINITION Gm Umb001.154 B19F UMN Soybean BAC Library (PESBAC4 BAC) Glycine

max genomic clone Gm Umb001.154\_B19; genomic survey sequence.

ACCESSION AQ989424

VERSION AQ989424.1 GI:6937370

KEYWORDS GSS.

SOURCE Glycine max (soybean)

ORGANISM Glycine max

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.  
1 (bases 1 to 532)  
AUTHORS Marek, L.F., Paz, M., Darnielle, L., Hanson, N. and Shoemaker, R. C.  
TITLE BAC End sequences from a soybean genomic library (ISU)  
JOURNAL Unpublished (2000)  
COMMENT Contact: Shoemaker Randy C  
Agronomy Department  
Iowa State University  
Ames, IA 50011-1010, USA  
Tel: 515 294 1205  
Fax: 515 294 2299  
Email: rcschoe@iastate.edu  
This BAC identified by SSR Sat429. For more information, see Soybase at:  
http://genome.cornell.edu/cgi-bin/WebPage/webase?db=soybase.  
Please see as an authority for the mapping/naming: Cregan P. B., T. Jarvik, A.L. Bush, R.C. Shoemaker, K.G. Lark, A.L. Kahler, N. Kaya, T.T. Vantol, D.G. Lohnes, J. Chung, and J.E. Specht. 1999a. An integrated genetic linkage map of the soybean genome. Crop Sci. 39: (in press)  
Seq primer: M13P  
Class: BAC ends.

## FEATURES

## SOURCE

## Location/Qualifiers

1..532  
/organism="Glycine max"  
/mol type="genomic DNA"  
/cultivar="Faribault"  
/db xref="taxon:3847"  
/clone="Gm-UMD001.154\_B19;"  
/tissue type="cotyledon leaves"  
/dev stage="cotyledon stage"  
/clone.lib="UMN Soybean BAC Library (PECSBAC4 Ecoti)"  
/note="Vector: pECSBAC4; The UMN BAC Library (Daneš et al, Theor. Appl. Genet. 96:196, 1998) was constructed using the Eco RI site of pECSBAC4. The library consists of 72,960 clones with an average insert size of 120 kb, equal to 7 haploid genome equivalents. Screening of the library is done by PCR amplification of DNA pools."

## ORIGIN

Query Match 41.2%; Score 227.8; DB 28; Length 532;  
Best Local Similarity 94.6%; Pred. No. 7.1e-51;  
Matches 229; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY 1 GTAGGTAATTCGAGACATTAACGACATTTGCGGAAGAAATGTAATAATAC 60  
DB 291 GTAGGTAATTCGAGACATTAACGACATTTGCGGAAGAAATGTAATAATAC 350  
QY 61 AAGTAGGACATAAATTACAAAACATGCGCTTTTGTCAAGATACGAATCTTCTC 120  
DB 351 AAGTAGGACATAAATTACAAAACATGCGCTTTTGTCAAGATACGAATCTTCTC 410  
QY 121 TTCTATCTATCTATCTATCTGAGTTCCTGTTGCAATGACGAACATTTCCGACA 180  
DB 411 TTCTATCTATCTATCTATCTGAGTTCCTGTTGCAATGACGAACATTTCCGACA 470  
QY 181 GGCGTAAATATAGCTCAAGTGGTGGGATCGTGTCTTGTATGAGATATTC 240  
DB 471 GGCGTAAATATAGCTCAAGTGGTGGGATCGTGTCTTGTATGAGATATTC 530  
QY 241 CT 242  
DB 531 CT 532

RESULT 14  
LOCUS BUS48744 712 bp mRNA linear EST 16-SEP-2002  
DEFINITION GM880015B20G10 Gm-r1088 Glycine max cDNA clone Gm-r1088-5708 3',  
ACCESSION BUS48744

VERSION  
KEYWORDS  
SOURCE  
ORGANISM

BUS48744.1 GI:22931605  
EST.  
Glycine max (soybean)  
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.  
1 (bases 1 to 712)  
AUTHORS Vokkin, L., Shoemaker, R., Keim, P., Retzel, E., Khanna, A., Shealy, R., Clough, S., Thibaud-Nissen, F., Corvill, V., Erpelting, J., Rapp, C., Shoop, E., Stromvik, M., Schweitzer, P., Gong, G. and Liu, L.  
TITLE A Functional Genomics Program for Soybean (NSF 9872565) (2002)  
JOURNAL Unpublished (2002)  
COMMENT Other ESTs: BF070750 corresponding to Gm-cl065-2303 (5')  
Contact: Vokkin, L.O., PI, A Functional Genomics Program for Soybean (NSF 9872565)  
Lewin, H. A., Director, Keck Center for Comparative and Functional Genomics  
University of Illinois  
Edwin R. Madigan Building, 1201 W. Gregory, Urbana, IL 61801, USA  
Tel: (217) 244-6147  
Fax: (217) 333-4582  
Email: l-vokkin@uiuc.edu  
Insert Length: 712 Std Error: 0.00  
Plate: GM880015B20 row: G column: 10  
Seq primer: 5'-TTTTTTTTTTTTTTTTT(A/C/G)-3'  
High quality sequence stop: 712.

## FEATURES

## SOURCE

## Location/Qualifiers

1..712  
/organism="Glycine max"  
/mol type="mRNA"  
/db xref="taxon:3847"  
/clone="Gm-r1088-5708"  
/clone.lib="Gm-r1088"  
/note="The library Gm-r1088 is a sequence-driven, rerecked set of 9,216 low redundancy clones selected from cDNA libraries from various tissues and stages of development of soybean that consists of 2,706 cDNAs from germinating cotyledons (source library Gm-cl027); 1,355 cDNAs from immature seed coats (libraries Gm-cl019, and Gm-cl023); 917 cDNAs from tissue culture derived somatic embryos (source libraries Gm-cl036 and Gm-cl075); 3,275 cDNAs from germinating seedlings, shoot tips, or leaves exposed to various stresses (source libraries Gm-cl064, Gm-cl065, Gm-cl066, and Gm-cl067); and Gm-cl068); and 963 cDNAs from young leaves exposed to bacterial and fungal pathogens (source libraries Gm-cl072, Gm-cl073, and Gm-cl074). The 5' ESTs of the source clones from the different libraries was used to select singletons, or a representative of each contig, which were rerecked to form library Gm-r1088 and the cDNA clones of the rerecked Gm-r1088 library were then sequenced at the 3' end. The unique selection and 3' sequencing was funded by NSF Plant Genome Project #9872565 (http://soybeanomics.croplsc.uiuc.edu/) as part of creation of a low redundancy soybean cDNA set. The source cDNA libraries were constructed by the laboratories of Lila Vokkin, University of Illinois, Randy Shoemaker, Iowa State University, and Paul Keim, Northern Arizona University as part of the Public EST Project, http://119.186.26.94/soybeanest.html. The contig analysis to select unique genes was performed by the laboratory of Ernest Retzel, Center for Computational Genomics and Bioinformatics, University of Minnesota.  
http://web.anc.uiuc.edu/biodata/nfssoy/. Rerecking and 3' sequencing were conducted by services of the University of Illinois Keck Center for Comparative and Functional Genomics http://www.biotech.uiuc.edu/keck.htm. Note: The corresponding 5' EST from each clone in the Gm-r1088 library is listed in the 'OTHER EST' field. The detailed information on the source library for each clone can also be obtained by referring to the Genome Systems clone ID of the original cDNA library that is also listed under 'OTHER

ORIGIN	EST'."
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

Query Match	40.5%;	Score	223.8;	DB	13;	Length	712;
Best Local Similarity	92.2%;	Pred. No.	8.8e-50;				
Matches	248;	Conservative	0;	Mismatches	17;	Indels	4;
						Gaps	1

QY	1	GTAGTATTATTCGACAGCACTTAACACAACTTTATTTTCGGGAAGAGTTGTACTAGAAATAC	60
Db	67	GTAAATTTTGGTCGACGACATTAACATTAATTATTTTCGGGAAGAGTTGTACTAGAAATAC	126
QY	61	AAGTAGGCACTAAAATTTCAAAATAATGCGATGGCTTTGTCCAGATACGAATCTTCTCTC	120
Db	127	AAGTACCGACCTCAAAATTACAAAATAATGCGATGGTTTGTCCAAATACGAAAATACTACTTC	186
QY	121	TTTCTATCTATCTATCTATCTCTGTGAGTCCGTGTGGAATTCGAGAACTTTTCCGAAACA	180
Db	187	TT-----TCATCTATCTATCTTGTAGATTCCTGTTCAAAATTGAGCAACTTTTCCGAAACA	242
QY	181	GGCTGTAAAAATAGCTTCAAGTCGTGGCGACTGTGTCTTGTCTTGAATGCAAGATATCC	240
Db	243	GGCTGTAAAAATAGCTTCAAGTCGTGGCAATCTGTGTCTTGTGAATGCAAGATATCC	302
QY	241	CTTAGCTTCACGACGTTCACCTTCCACCTCG	269
Db	303	CTTAGCTTCACGACGTTCACCTTCCACCTCG 331	

RESULT 15	LOCUS	DEFINITION	LOCUS	DEFINITION
BG586953/c	BG586953	726 bp mRNA linear EST 11-APR-2001	BG586953	726 bp mRNA linear EST 11-APR-2001
	EST488722	MHAM Medicago truncatula/Glomus versiforme mixed EST	EST488722	MHAM Medicago truncatula/Glomus versiforme mixed EST
	library	CDNA clone pMHAM-47E20 5' end, mRNA sequence.	library	CDNA clone pMHAM-47E20 5' end, mRNA sequence.

directionally ligated into the Unizap XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-Zap phage using Ex-assist helper phage and propagated in XL10R cells."

Query Match	40.0%	Score 221.2	DB 12	length 726
Best Local Similarity	85.2%	Pred. NC. 4.5e-49		
Matches 247	Conservative	0	Mismatches 43	Indels 0
				Gaps 0

QY	139	TCTGTGAGTTCCTTTGGAAATGAGCAACCTTTCCGAAACAGCTGTAAATAATAGCTTC	198
Db	661	TCTTTGAGTTCCTTTGCAATGAGCAACCTTTCCGAAACAGCTGTAAATAATAGCTTC	602
QY	199	AAGTGTTGGCAATCTGTGCTCTTTGATGCAAGATATCCCTTGCTTCACTGAGCTGC	258
Db	601	AAGTGTTGGCAATTTTCAATTTTATCTTGATGAGAGATATCCGATGCTCTTAAAGTCTGC	542
QY	259	ACTTCGACCTCGCAAGTCATCATGATGCCAAATGAGAACTGTGGACATCATATATGAGCC	318
Db	541	ACTTCGATCTCGAAGTATCATGATGCCAAATGAGATCTTCGGAAACATGATATGCCCC	482
QY	319	CCAAATAACTTCTCTCTCTGTATGATCATTCGCCCTTGACTTTTAAAGTCAACTCA	378
Db	481	CCAAGCAACTTCTTATTTCTGTATTAACCATTCCTCTTGACTTTTAAAGCAACATCA	422
QY	379	TATGTATPACACAAGCTTCTTGTTCCGAAGTATGACCAATGAAATGCAAC	428
Db	421	TATGTAAACTGACACGTTTCTTGTTCCGAACGCACTACAAAGTGAATC	372

Search completed: March 19, 2004, 21:33:14  
Job time : 2135 secs

[illegible]

```

Db      1378 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYGGTA 1437
Cy      453 CGAGTACTGTGAAAAATTAGCT 476
        |||||
Db      1438 CCAATTCTCTACTCTTAACT 1461

RESULT 3
US-09-539-333D-129/c
: Sequence 129, Application US/09539333D
Patent No. 6476208
: GENERAL INFORMATION:
: APPLICANT: Cohen, Daniel
: APPLICANT: Blumenfeld, Marta
: APPLICANT: Chumakov, Ilya
: APPLICANT: Bougueleret, Lydie
: APPLICANT: Bihain, Bernard
: APPLICANT: Esbloux, Laurent
: TITLE OF INVENTION: SCHIZOPHRENIA ASSOCIATED GENES, PROTEINS AND BIALLELIC MARKERS
: FILE REFERENCE: GENSET .047AUS
CURRENT APPLICATION NUMBER: US/09/539,333D
CURRENT FILING DATE: 2000-03-30
PRIORITY FILING DATE: 2000-03-30
PRIORITY FILING DATE: 1999-03-30
PRIORITY FILING DATE: 1999-03-30
PRIORITY FILING DATE: 1999-04-30
PRIORITY FILING DATE: 1999-04-30
PRIORITY FILING DATE: 1999-04-30
PRIORITY FILING DATE: 1999-04-30
PRIORITY FILING DATE: 1999-07-14
PRIORITY FILING DATE: 1999-07-27
PRIORITY FILING DATE: 1999-07-29
PRIORITY FILING DATE: 1999-07-29
PRIORITY FILING DATE: 1999-07-29
PRIORITY FILING DATE: 1999-10-28
PRIORITY FILING DATE: 1999-10-28
PRIORITY FILING DATE: 1999-10-12
NUMBER OF SEQ ID NOS: 231
SOFTWARE: Patent.pm
SEQ ID NO 129
LENGTH: 3001
TYPE: DNA
ORGANISM: Homo Sapiens
FEATURE:
NAME/KEY: allele
LOCATION: 1501
OTHER INFORMATION: 99-25966-241 : polymorphic base T or C
FEATURE:
NAME/KEY: misc_binding
LOCATION: 1502..1520
OTHER INFORMATION: 99-25966-241.misl, complement
FEATURE:
NAME/KEY: misc_binding
LOCATION: 1481..1500
OTHER INFORMATION: 99-25966-241.mis2,
FEATURE:
NAME/KEY: primer_bind
LOCATION: 1721..1741
OTHER INFORMATION: upstream amplification primer, complement
FEATURE:
NAME/KEY: primer_bind
LOCATION: 1219..1239
OTHER INFORMATION: downstream amplification primer
FEATURE:
NAME/KEY: misc_binding
LOCATION: 1489..1513
OTHER INFORMATION: 99-25966-241 probe
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1659,1987

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 19, 2004, 20:20:31 ; Search time 72 Seconds  
(without alignments)  
4262.331 Million cell updates/sec

Title: US-09-531-113-5981

Perfect score: 553  
Sequence: 1 gtaggttattccagacatt.....ccaaggagactgacgag 553

Scoring table: IDENTITY NUC  
Gapop 10.0, Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Issued Patents NA: \*  
1: /cgn2\_6/prodata/2/ina/5A COMB.seq: \*  
2: /cgn2\_6/prodata/2/ina/5B COMB.seq: \*  
3: /cgn2\_6/prodata/2/ina/6A COMB.seq: \*  
4: /cgn2\_6/prodata/2/ina/6B COMB.seq: \*  
5: /cgn2\_6/prodata/2/ina/PCITUS COMB.seq: \*  
6: /cgn2\_6/prodata/2/ina/backfile1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	98.6	17.8	277	4	US-09-313-294A-1048
C 2	53.2	9.6	7218	1	US-08-232-463-14
C 3	35.8	6.5	3001	4	US-09-539-333D-129
C 4	34.4	6.2	58073	4	US-08-545-528D-1
C 5	33.8	6.1	1500	4	US-09-134-001C-2533
C 6	33.6	6.0	246	4	US-09-280-116-144
C 7	33.4	6.0	193303	4	US-09-497-855A-37
C 8	33.4	6.0	193303	4	US-09-497-855A-44
C 9	33.3	6.0	399	4	US-09-621-976-8976
C 10	33.3	6.0	2135	4	US-08-933-711B-17
C 11	33.3	6.0	4612	3	US-08-960-048-2
C 12	33.3	6.0	4612	3	US-08-938-586-2
C 13	33.3	6.0	55827	4	US-09-813-133A-3
C 14	33.3	6.0	640681	4	US-09-790-998A-1
C 15	33.3	6.0	1664976	4	US-08-916-421B-1
C 16	32.8	5.9	1751	4	US-09-453-702B-223
C 17	32.2	5.8	2169	4	US-09-434-408-3
C 18	32.2	5.8	2451	4	US-09-134-000C-2752
C 19	32.2	5.8	3336	4	US-09-717-364A-26
C 20	32.2	5.8	7291	3	US-08-913-943-3
C 21	32.2	5.8	26270	4	US-09-717-364A-1
C 22	31.8	5.8	1996	3	US-08-946-026-15
C 23	31.8	5.8	6124	4	US-08-213-419B-3
C 24	31.8	5.8	6773	4	US-09-166-350-27
C 25	31.6	5.7	1155	3	US-08-860-368B-19
C 26	31.4	5.7	966	4	US-09-543-681A-3120
C 27	31.2	5.6	3101	4	US-09-221-017B-87

C 28	31.2	5.6	3455	4	US-09-651-656-20	Sequence 20, App1
C 29	31.2	5.6	3455	4	US-09-650-855-20	Sequence 20, App1
C 30	31	5.6	873	2	US-08-531-927B-6	Sequence 6, App1
C 31	31	5.6	876	3	US-09-315-793-61	Sequence 61, App1
C 32	31	5.6	1512	3	US-08-945-994-4	Sequence 4, App1
C 33	31	5.6	1776	2	US-08-531-927B-1	Sequence 1, App1
C 34	31	5.6	1776	3	US-09-041-886-12	Sequence 12, App1
C 35	31	5.6	4176	4	US-09-722-139-1	Sequence 1, App1
C 36	31	5.6	4176	3	US-09-721-832-1	Sequence 1, App1
C 37	31	5.6	4176	4	US-09-721-832-1	Sequence 1, App1
C 38	31	5.6	6254	4	US-08-956-171E-404	Sequence 404, App
C 39	31	5.6	7015	3	US-09-177-249-6	Sequence 6, App1
C 40	31	5.6	14066	4	US-09-601-198-56	Sequence 56, App1
C 41	31	5.6	1664976	4	US-08-916-421B-1	Sequence 1, App1
C 42	31	5.6	1830121	4	US-09-557-884-1	Sequence 1, App1
C 43	31	5.6	1830121	4	US-09-643-990A-1	Sequence 1, App1
C 44	30.8	5.6	6617	4	US-09-976-594-268	Sequence 268, App
C 45	30.6	5.5	1047	4	US-09-543-681A-693	Sequence 693, App

## ALIGNMENTS

RESULT 1  
US-09-313-294A-1048/C  
Sequence 1048, Application US/09313294A  
Patent No. 6476212  
GENERAL INFORMATION:  
APPLICANT: Ialoudi, Raghunath V.  
APPLICANT: Ito, Laura Y.  
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR  
FILE REFERENCE: PL-0017 US  
CURRENT APPLICATION NUMBER: US/09/313,294A  
CURRENT FILING DATE: 1999-05-14  
NUMBER OF SEQ ID NOS: 7600  
SOFTWARE: PERL Program  
SEQ ID NO 1048  
LENGTH: 277  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:  
NAME/KEY: misc. feature  
OTHER INFORMATION: Incyte ID No. 6476212 700550303H1  
US-09-313-294A-1048

Query Match 17.8%; Score 98.6; DB 4; Length 277;  
Best Local Similarity 61.5%; Pred. No. 1.9e-19;  
Matches 158; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY	167	ACTTTCCCGACAGGCTGTAAATAAGCTCAAGTCGTGGGCACTGTGCTTCTT	226
DB	257	ATTTCCTCTGATGGGCGCCAAAGACTTGAGATCTTGAGATCCGATCTTCAAT	198
QY	227	GATGCAAGATATCCCTTGTCTTCACTGCTTCCACCTGCAAGTATGATTTAC	286
DB	197	CAGAGCGAGGCGCTTGATCGCGAAGCTTACTTGACCTTCAAGTATGATTTAC	138
QY	287	CAAGGAGACTCTGGACATTTATATGACCCCAATAAATCTTCTTCTTGATTA	346
DB	137	CGAAGGAAATCGGGAATGTCTAATGCGCTTGTATCTTGTCTTCTTGATTA	78
QY	347	TCCATTCCTTGAAGTCACTATGATATATACACAGTTCTTCTTCTTCC	406
DB	77	ACCATTCACCGTAATCTCAAGCTTATGATTAATCTTCTTCTTCTTCTTCTT	18
QY	407	GAATATCACTCAAT 423	
DB	17	GGCAATGACCTGAAT 1	

RESULT 2  
US-08-232-463-14



OTHER INFORMATION: n=a, g, c o r t

SEQ ID NO. 129

Query Match 6.5%; Score 35.8; DB 4; Length 3001;  
Best Local Similarity 54.1%; Pred. No. 0.9;  
Matches 73; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 419 TGAATGCAACATTTTCTTAATTAATCTGAACCTGACGATCCTGAAAAATTAGCTGA 478  
DB 2385 TGCAATATCATTTTGGCCAGCTGGAAGCTAGAGAGTCAAGAAAAACAACAA 2326  
QY 479 AAAAAACCATTTCTATCGCAACGATTTGTTACATCTTTACTTTGCTCCGCAAG 538  
DB 2325 AAAACCTGGAATTAACAGCTGACTAGGCTTGATCAATTAATGACCTTGGCCCTGACAT 2266  
QY 539 GAGACTGGAATGAG 553  
DB 2265 GAGTATGATTTGAG 2251

## RESULT 4

US-08-545-528D-1  
; Sequence 1, Application US/08545528D  
; Patent No. 6537773  
; GENERAL INFORMATION:  
; APPLICANT: Fraser et al.  
; TITLE OF INVENTION: Nucleotide Sequence of the Mycoplasma Genitalium Genome, Fragment  
; Patent No. 6537773  
; TITLE OF INVENTION: Thereof, and Uses Thereof  
; FILE REFERENCE: PB193P1  
; CURRENT APPLICATION NUMBER: US/08/545,528D  
; PRIOR FILING DATE: 1995-10-19  
; PRIOR APPLICATION NUMBER: US 08/488,018  
; PRIOR FILING DATE: 1995-06-07  
; PRIOR APPLICATION NUMBER: US 08/473,545  
; NUMBER OF SEQ ID NOS: 1  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 1  
; LENGTH: 580073  
; TYPE: DNA  
; ORGANISM: Mycoplasma genitalium  
US-08-545-528D-1

Query Match 6.2%; Score 34.4; DB 4; Length 580073;  
Best Local Similarity 48.9%; Pred. No. 17;  
Matches 92; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 366 TAAGTCACTCATATGATTAACAACAACGTTCTTCTGCGACTATCACCATGATGC 425  
DB 281067 TGAGATAAACAATGATTAACAACAACAACCAATTTTAAACCACTTGATGGCTTAAGC 281126  
QY 426 AACATTTCCATATATCTGAACCTGAGAGTACTTGAATAAATTAGCTGAAAAAAC 485  
DB 281127 AACATTTTCTGATTAATATCTACCAAAAAATCTTTAAGAAATATCTATCAGAAAAATC 281186  
QY 486 CTGATTCATGCGCCACGATTTTGTATCATCTTTTACTCTGCGCCGCAAGAGACT 545  
DB 281187 AACAGCACTTAACAATTTCTTTTAAACAGTTTGAAGAAATGATTTGTTGCTCAAAA 281246  
QY 546 GAATGAG 553  
DB 281247 AAAAGAG 281254

## RESULT 5

US-09-134-001C-2533  
; Sequence 2533, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: GTC-007

; CURRENT APPLICATION NUMBER: US/09/134,001C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 2533  
; LENGTH: 1500  
; TYPE: DNA  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-2533

Query Match 6.1%; Score 33.8; DB 4; Length 1500;  
Best Local Similarity 54.4%; Pred. No. 2.6;  
Matches 68; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

QY 102 CAGATACGAATCTCTCTTCTATCTATCTATCTATCTGAGTCTGCTGGAATTG 161  
DB 146 CAGGCGCTATCTCTCTTCTATCTATCTATCTATCTGAGTCTGCTGGAATTG 205  
QY 162 CAGCACTTTTCCGAGACAGGCTGAAAAATAGCTTCAAGTCTGCGCATCTGTCTT 221  
DB 206 CTGTAAGTATGCGCCACTGCTTGAATAATTTCTGTTCTGCTGCGCATTTGCTTTA 265  
QY 222 GTCTT 226  
DB 266 GTTTT 270

## RESULT 6

US-09-280-116-144  
; Sequence 144, Application US/09280116A  
; Patent No. 6331427  
; GENERAL INFORMATION:  
; APPLICANT: Robison, Keith E.  
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs  
; FILE REFERENCE: 5800-24, 035800/175965  
; CURRENT APPLICATION NUMBER: US/09/280,116A  
; PRIOR FILING DATE: 1999-03-26  
; NUMBER OF SEQ ID NOS: 268  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 144  
; LENGTH: 246  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: trypsin-like serine proteases  
US-09-280-116-144

Query Match 6.1%; Score 33.6; DB 4; Length 246;  
Best Local Similarity 52.9%; Pred. No. 1.5;  
Matches 72; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 286 CCAATGAGACTCTGGGACATCTATATGACCCCAATAAATCTTCTCTCTTGATG 345  
DB 38 CCAATGAGAGCAAGGCTCTCCCAATGACCAAGAAATGATGATGCTTCCCTCAAG 97  
QY 346 ATCCATTCCTCTTGAATTTAAGTCACTCATATGATATACCAACAGCTTCTTCTT 405  
DB 98 CTGCTCCAGAGTTCATTTTCTTATCTCTATCTGATGAGAGGCTGACACTGATC 157  
QY 406 CGAATATCACCATGA 421  
DB 158 CGGCTGCAATCARGA 173

## RESULT 7

US-09-497-855A-37  
; Sequence 37, Application US/09497855A  
; Patent No. 6605432  
; GENERAL INFORMATION: